THE UNIVERSITY OF ALBERTA

THE DEVELOPMENT OF ATTRIBUTIONS OF RESPONSIBILITY

by

© KENNETH ANTHONY HILL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

Spring, 1977

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "The Development of Attributions of Responsibility", by Kenneth Anthony Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

.....
Supervisor

.....

.....

.....

.....
External Examiner

Date.....November 22, 1976.....

Dedication

For Chandra, who is teaching me about childhood.

Abstract

This dissertation investigated the development of children's attributions of responsibility for both good and bad consequences. It was proposed that there is research evidence indicating that persons attribute more personal responsibility to individuals involved in acts leading to bad consequences than good ones. However, different explanations were considered necessary when young children rather than older children and adults demonstrate such a valence-of-consequence effect. Hypotheses derived from the relevant attribution theory literature were proposed in the case of older children, whereas the moral judgment literature was the main source of hypotheses pertaining to young children. Moreover, it was proposed that information regarding parental dictates and peer group consensus would have differential and age-related effects upon children's attributions of responsibility.

Two experiments were conducted with boys and girls in grades two, five and eight. In both experiments, subjects heard stories of fictional children involved in acts leading to good or bad consequemces of high or low severity, and were subsequently required to indicate attributions of responsibility and subsidiary attributions on four-point scales. In Experiment 1, actors were said to have performed such actions in accordance with parental dictates (i.e., the dictate was facilitative) or despite the parental dictates (i.e., the dictate was inhibitory), or no information as to parental dictate was given (control). Results

Abstract

This dissertation investigated the development of children's attributions of responsibility for both good and bad consequences. It was proposed that there is research evidence indicating that persons attribute more personal responsibility to individuals involved in acts leading to bad consequences than good ones. However, different explanations were considered necessary when young children rather than older children and adults demonstrate such a valence-of-consequence effect. Hypotheses derived from the relevant attribution theory literature were proposed in the case of older children, whereas the moral judgment literature was the main source of hypotheses pertaining to young children. Moreover, it was proposed that information regarding parental dictates and peer group consensus would have differential and age-related effects upon children's attributions of responsibility.

Two experiments were conducted with boys and girls in grades two, five and eight. In both experiments, subjects heard stories of fictional children involved in acts leading to good or bad consequences of high or low severity, and were subsequently required to indicate attributions of responsibility and subsidiary attributions on four-point scales. In Experiment 1, actors were said to have performed such actions in accordance with parental dictates (i.e., the dictate was facilitative) or despite the parental dictates (i.e., the dictate was inhibitory), or no information as to parental dictate was given (control). Results

indicated that control-group children of all grades attributed more responsibility when consequences were bad rather than good ($p < .05$). Other major findings include a significant tendency for children of all grades to attribute comparable levels of responsibility when the parental dictate was facilitative to a good consequence and when no such information was given, with increased responsibility attributed when the dictate was facilitative. This interaction was significant at the .01 level. Also supported was the prediction that second-grade children--and only second-grade children--would judge actors to be less good ($p < .05$) and allocate less reward ($p < .05$) to actors who performed good acts despite parental dictates to the contrary, compared to actors who complied with facilitative parental dictates in performing the same good acts.

Experiment 2 involved similar independent and dependent variables, except that peer consensus rather than parental dictates served as the external cause manipulation. While the valence main effect of Experiment 1 was replicated ($p < .01$), it was found that consensus information in general did not have a potent effect upon the dependent variables for any grade level. However, there was evidence that children would attribute more responsibility to actors involved in good consequences when peer consensus was inhibitory to performance of the good act than when it was facilitative ($p < .05$). Predictions regarding age differences on this interaction were not supported.

The general discussion focused upon the moral judgment-responsibility attribution relationship, the role of peer consensus as an influence upon behavior vs. an influence upon perceptions of social causality, and possible cognitive developmental factors to be considered in attribution theory.

Acknowledgements

I am extremely grateful to Brendan Gail Rule for her excellent supervision and considerate guidance over these long years, and to Michael Enzle for his efforts as teacher, colleague and friend. I also wish to thank Birenda Sinha and Charles Hobart for their service as committee members, and Willard Hartup for serving as an external examiner. In addition, I wish to thank Merrilyn Greig and Natalie Daviduk for graciously volunteering their services as typists. Finally, I am thankful for Carol, who helped with the pilot study and endured many lonely evenings while this thesis was being written.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The purpose of this dissertation is to investigate age-related changes in attributions of responsibility. Of particular concern is the frequently observed tendency for both adults and children to attribute more personal responsibility to actors involved in negative outcomes than in positive outcomes. However, it is proposed that this effect, referred to as the valence-of-outcome effect, occurs for different reasons in young and older children. It is suggested that current attribution-theoretical terminology can explain the phenomemon in older children, whereas specific aspects of the parent-child relationship account for the valence-of-outcome effect in young children. Relevant theoretical and empirical contributions in the moral judgment literature are discussed with regard to the latter explanation. The results of two experiments aimed toward testing hypotheses derived from the present discussion are presented. Finally, implications of experimental results for the understanding of the development of attribution processes, and for attribution theory in general, are discussed.

Attributions of Responsibility and the Valence-of-Outcome Effect

Heider's (1958) original formulations on attributions of responsibility spurred considerable research effort in that area. Researchers have been particularly concerned with situational factors influencing the degree to which persons judge actors responsible for positive as well as negative consequences.

Of particular interest is the frequent observation that the valence of an act's outcome (i.e., good or bad) seems to be an important factor in attributions of personal responsibility by both children and adults (Schneider & Shaw, 1970; J. Shaw & Skolnick, 1971; M. Shaw & Reitan, 1969; M. Shaw & Schneider, 1969; M. Shaw & Sulzer, 1964; Weiner & Peter, 1973). In general, more personal responsibility seems to be attributed to persons involved in acts leading to negative rather than positive outcomes, and the former are typically allotted more punishment than the latter are rewarded. Although developmental trends have not been directly tested, this valence-of-outcome effect appears to be particularly evident with younger children and to diminish somewhat with age (cf. Schneider & Shaw, 1970; Shaw & Sulzer, 1964; Weiner & Peter, 1973).

Although Kelley's extension of attribution theory (1967, 1971a) has been applied to attributions of responsibility as well (e.g., Kelley, 1971b; Ross & DiTecco, 1975), attribution theorists have not yet provided a systematic explanation for the valence-of-outcome effect. Kelley (1971b) suggested the possibility that, when attributions of responsibility are being made, plausible external causes may be more likely to have a mitigating role in good outcomes than bad. Kelley emphasized the role of consensus information as an important determinant of internal versus external attributions of responsibility.[1] He proposed that if there is high consensus (i.e., everyone does it), internal attributions will be low. On the other hand, perceived low

consensus should lead to increased attributions of responsibility to the person. According to Kelley (1971b), the problem with assigning internal responsibility to an actor for his good acts is that "goodness" is based on consensus, and is therefore externally caused. On the other hand, when acts have negative outcomes, low consensus is assumed and high internal responsibility is therefore attributed. Ross and DiTecco (1975) have similarly emphasized the role of consensus information in attributions of responsibility, and have proposed that persons are not judged responsible for transgressions which are also committed by a great number of people.

Moreover, although it is implied in Kelley's discussion of moral evaluations and attributions of responsibility (1971b), Kelley has not applied his more specific hypothesized functions of external causes to attributions of responsibility, i.e., whether the cause is perceived to be facilitative or inhibitory (Kelley, 1971a). An external cause that is facilitative is one that can be assumed to have plausibly caused an observed effect, either solely or in conjunction with causes internal to the actor. An external inhibitory cause is one that can be assumed to have militated against the occurrence of an effect, as when a strong current thwarts the upstream progress of a rowboat. An application of this distinction might provide a more systematic and comprehensive explanation for the valence-of-outcome effect than is suggested in Kelley's (1971b) lucid but rather cursory treatment of attributions of responsibility.

A plausible facilitative external cause is presumed to lead to decreased attributions of personal responsibility, a process conforming to what Kelley termed the discounting principle. On the other hand, the augmenting principle refers to when there exists a plausible external inhibitory cause, i.e., one which is in opposition to the observed outcome, which should therefore lead to attributions of higher internal responsibility than when no such external cause is present. Kelley's (1971b) discussion of moral evaluations leads to the hypothesis that individuals will assume facilitative external causes (i.e., consensus information) when outcomes are good, even when no such information is explicitly given. However, an inhibitory external cause should be assumed when consequences are bad. As the discounting and augmenting principles would therefore be differentially invoked, depending upon whether the outcomes are positive or negative, respectively, the facilitative-inhibitory distinction thus has direct bearing upon the valence-of-outcome effect.

However, it is unlikely that, where young children are concerned, the valence-of-outcome effect can be explained according to such attributional biases, since studies with children on attributions of social causality (Smith, 1975) and physical causality (Shultz et al., 1975) indicate that children under nine years of age do not process external-cause information well. It does not seem likely, therefore, that external factors such as perceived consensus could have mitigating effects on young children's attributions of responsibility. Thus, it would appear

that, although an attributional analysis may be appropriate to an explanation of the valence-of-outcome effect in older children and adults, a different explanation may be necessary in the case of young children. It is the moral judgment literature in general and Piaget (1932) in particular where this explanation will be sought.

Moral Judgments and the Valence-of-Outcome Effect

In his theory of moral development, Piaget (1932) proposed that the young child's basis for moral evaluations is not peer consensus, but rather the child's unqualified respect for his parent's authority, which constitutes his sole source of his understanding of right and wrong. Because the young child's morality was presumed to be externally imposed upon him, and therefore as "objective" as physical reality, he was said to respond rigidly and punitively to actions which contradicted what he perceived as morally "right." However, as the child becomes more experienced in social interaction with his peers, he is described by Piaget as learning to appreciate the "subjective" nature of moral obligation. During this second stage, termed the morality of cooperation, moral rules serve to reconcile the differing expectations and points of view of the child's peer group, thereby becoming more functional to optimal social interaction.

While this particular aspect of Piaget's theory has been criticized for its conceptual ambiguities and has not held up

well under empirical investigation (cf. Kohlberg, 1969), it should be noted that he was justifiably impressed by the child's relatively abrupt attachment to the peer group, as this expanded social focus is undoubtedly a crucial step in the socialization process (cf. Hartup, 1970). For our purposes, Piaget's theory suggests a possible addendum to attributions of responsibility, at least where young children are concerned. While we may be correct in describing the older child as making his attributions of responsibility following a causal analysis, especially a consideration of consensus information, Piaget's theory and the discussion in the previous section suggest that younger children may not only fail to make such a causal analysis, but to attribute responsibility to specific parental dictates rather than peer consensus.

Relevant to the valence-of-outcome effect is Piaget's (1932) proposition that young children, because of their response to parental authority, have considerably more difficulty in understanding good than bad behavior. He suggested that the young child's concept of good is based on duty or obligation, rather than the more autonomous, non-obligatory concept of good understood by older children and adults. Similarly, Hill and Enzle (Note 1) suggested that parents are typically inclined to focus the young child's moral education around the concept of naughtiness and the avoidance thereof, while the child's understanding of good behavior is confused with compliance with parental dictates. Baldwin et al. (1971), in a cross-cultural study,

explained their results in a similar fashion. In that study, a substantial number of young children judged an obedient child as "kinder" than a spontaneously helpful one. This was particularly evident for children in Yucatan, where, the authors maintained, obedience is an intensely socialized value.

Thus, it may be that it is the nature of child-rearing practices and not causal inferences which is the cause of the valence-of-outcome effect in young children. More specifically, it may be the young child's differing and specific reactions to good and bad acts which produce the effect. For example, several research reports indicate that young children are considerably influenced by the consequences of an act when making judgments of good to an actor, while they are much less influenced by consequences when judging bad behavior (Hill & Enzle, Note 1; Jenson & Hughston, 1973; McKechnie, 1971). Jenson and Hughston (1973) varied good, neutral and bad content acts followed by reward or punishment, and found that young children were very much inclined to judge good and neutral content acts according to sanctions and not intentions, whereas bad content acts were usually judged to be bad regardless of sanctions. Thus, it is possible that the young child may not perceive the inherent "goodness" of another person's behavior before he has grasped the more mature, non-obligatory concept held by the older child and adult. He may therefore fail to understand the concept of intending or trying to commit a good act, and thereby fail to attribute responsibility for such acts, especially when the out-

comes are relatively mild. On the other hand, as the Jenson and Hughston (1973) study illustrates, children seem to be very knowledgeable in regard to the meaning of bad behavior. A recent study by Berg-Cross (1975) supports this suggestion. Her first-grade subjects judged bad acts resulting from obedience to authority as nevertheless deserving of punishment. These findings are reminiscent of Piaget's (1932) description of the young child as a harsh and unforgiving judge of transgressors.

This discussion therefore proposes that the valence-of-outcome effect in young children is a dual function of their failing to consider good intentions (and thus failing to assign responsibility accordingly) and an over-reaction to bad behavior (and thus elevating their attributions of responsibility). Moreover, the discussion above also leads to the prediction that extenuating factors should have little or no effect upon the attributions of responsibility by young children. This is the case because such children are not assumed to be making causal inferences but rather responding rigidly to specific characteristics of situations.

Statement of the Problem

A complaint that is sometimes heard is that we are reprimanded more for our vices than rewarded for our virtues. While this seems to be supported by research evidence, little more than passing explanatory consideration has been given to it. Kelley (1971b) alone among recent psychologists has discussed the irony.

The problem, Kelley suggested, lies in the fact that prosocial behavior is normally assumed to be non-extraordinary, i.e., that "anyone" would act prosocially in appropriate situations, and thus it is the situation and not the person which is responsible for the behavior. Conversely, bad acts do implicate the perception of uniqueness and therefore are seen as caused by actors involved. These propositions have received only very indirect support (e.g., Cunningham & Kelley, 1975) and have not yet been directly investigated.

Moreover, an investigation of the manner in which perceivers weigh the relative contributions of persons versus situations when making attributions of responsibility would require specifying the facilitative versus inhibitory nature of the situational cause. For example, a prosocial act performed in the face of consensus information at variance with the act's performance may actually augment responsibility attributed to the actor, compared to the same act performed in the absence of consensus information. The possibility that the nature of external causes may contribute to the valence-of-outcome effect remains untested, although it is implied in Kelley's theory, discussed above. For example, Kelley suggests that perceivers assume facilitative external causes when attributing responsibility for good consequences, even when no such information is given directly. Conversely, inhibitory external causes may be similarly assumed in the case of bad outcomes. These differences alone should lead to differences in perceived trying on the part of actors, resulting in

more internal responsibility attributed for those involved in negative consequences. However, as stated above, such attributional biases are yet to be empirically demonstrated.

On the other hand, an attribution-theoretical explanation for the valence-of-outcome effect, with its emphasis upon rather complex cognitive processes, is unlikely to be appropriate in accounting for the effect in young children, especially in light of research indicating an absence of relatively sophisticated causal analyses im children under nine years of age. Research in the moral judgment literature indicates that it is the young child's unique understanding of the concept of good (judgments of which may be affected more by the degree of consequences than intent) and his relatively punitive orientation to bad behavior which may be causing the effect, rather than attributional biases. Furthermore, this literature indicates that such judgments by young children are influenced primarily by parental authority rather than the nature of consensus information. However, these suggestions have never been directly examined in the context of attributions of responsibility, and their pertinence to an explanation for the valence-of-outcome effect in young children therefore remains speculative.

A final issue concerns the developmental pattern of attributions of responsibility. Practically nothing is known regarding the manner in which the young child gradually acquires a more sophisticated grasp of social causality, to where he is perform-

ing the relatively complex causal analyses demonstrated in later childhood. For example, it would be instructive to assess the approximate age at which the child begins to exhibit the discounting and augmenting principles and when and how such attributional phenomena are affected by the valence of outcomes. Such information bears upon the general area of attribution theory as well as the specific issues addressed herein.

General Design and Hypotheses

The general procedure involved subjects rendering judgments regarding the behavior of story characters involved in good and bad consequences. Two studies were conducted. In Experiment I, there were three grades (second, fifth and eighth), two outcome valences (good and bad), two levels of consequences (high and low), and three types of information regarding external causes (inhibitory, facilitative, no information). For purposes of expediency, external cause was a between-subjects factor, while outcome valence and outcome severity were crossed within subjects. The external cause manipulation for Experiment 1 consisted of parental dictates. Thus, a subject in this experiment had to make judgments regarding story characters involved in high- and low-severe bad and good acts. Further, story characters had been ordered to do or not to do all such acts, or the acts were performed with no information regarding dictates being given. Therefore, each subject was exposed to a total of four stories. All acts were performed intentionally by

story characters.

In Experiment 2, grades (second, fifth and eighth), outcome valence (good and bad) and severity of outcomes (high and low) were again varied. Also, external cause information (inhibitory, facilitative, no information) was varied, but for Experiment 2 this information involved peer consensus rather than parental dictates. Thus, a subject in this experiment was exposed to four stories concerning story characters who performed acts which were in accord with or counter to peer consensus, or no information regarding consensus was given.

Both experiments contained the same three major dependent variables: attributed responsibility, sanctioning, and diagnosticity. Responsibility was defined explicitly to the child (see Method section). The sanctioning measure consisted of asking the child to rate the degree to which the actor should be rewarded or punished for the act, depending on the valence of the outcome. Finally, diagnosticity (cf. Brickman, Ryan & Wortman, 1975) was measured by asking the child how much he or she would expect the actor to perform similar acts in the future. Diagnosticity follows directly from Kelley's hypothesis that external attributions generalize across persons and internal attributions across situations (cf. Cunningham & Kelley, 1975). This distinction has also been made by Brickman, Ryan & Wortman, (1975).

These three dependent variables were included not because differential predictions were made, but in order to explore the

relationships between these particular variables. For example, Shaw and Reitan (1969) reported that, while attributions of responsibility and sanctioning were highly and positively correlated when outcomes were negative, subjects were less inclined to reward persons for good acts for which they had been held responsible. Similarly, while diagnosticity is an important concept in attribution theory and has received some empirical consideration in the context of trust and surveillance (e.g., Strickland, 1958), more information should be sought regarding diagnosticity, independently of other attributions. Also, it was considered worthwhile to measure developmental trends in diagnosticity, especially as it related to the other dependent variables.

Generally, main effects were expected for each of the four factors (valence, severity, grades, and inhibitory vs. facilitative cause) on all dependent measures, in both the dictate and consensus studies. However, as would follow from the previous discussion, significant interactions were also expected. These more specific hypotheses are as follows:

Experiment 1 (Dictate). First, it was expected that grades would interact with valence, with the valence-of-outcome effect being more salient with second-grade children than for fifth- and eighth-grade children. Grades was also expected to interact with severity of outcomes, with second-graders attributing more responsibility when outcomes were relatively severe than would fifth-graders, who in turn were expected to be more

influenced by severity than would eighth-grade subjects. Finally, grades was expected to interact with cause (inhibitory vs. facilitative). While the attributions of second-grade children were expected to be affected by both valence and severity of outcomes, it was not expected that they would be influenced by information regarding the facilitative nature of the external cause in the same manner as would eighth-grade children. In particular, in the inhibitory cause/ good outcome condition (i.e., where the child is asked to judge an actor who has helped someone despite orders to the contrary from an adult), it was considered likely that many of the second-graders would tend to see the actor as performing an act of disobedience and would judge him harshly. This would follow if, as maintained previously, the young child's concept of good is confused with obedience to parents. However, whem external causes are facilitative (i.e., the child performs an act accordantly with a parental dictate) or no such information is given, the second-graders should tend to attribute responsibility and to allocate rewards according to the severity of the outcomes, with the facilitative-cause situation probably yielding the highest reward. In this case, we cannot say that the child is discounting or augmenting internal responsibility, as would be predicted from attribution theory for the older children, but rather is responding to specific aspects of situations with little systematic consideration of causal information.

This lack of regard for the causal influence of mitigating

factors (in this case, parental dictates) is further illustrated in the bad-outcome situation. In this case, however, moral judgment research (e.g., Berg-Cross, 1975; Jenson & Hughston, 1973) indicates that young children may not be affected by the obedience-disobedience information (i.e., information which serves as inhibitory and facilitative causes for older children), and degrees of sanctioning and attributed responsibility are expected to be influenced primarily by severity of outcome.

On the other hand, eighth-graders are expected to be less influenced by the valence and severity of outcomes than by the facilitative or inhibitory nature of the external cause preceding the act (although this hypothesis will be qualified shortly, when differences between dictates and consensus are considered). Fifth-grade children are expected to occupy a relatively intermediate position, although the literature supports the possibility that they may respond more like eighth-graders than like second-graders.

Experiment 2 (Consensus). As in the dictate study, age is expected to interact with valence and also with severity of outcomes. Moreover, age is expected to interact with the facilitative-inhibitory nature of external cause, but not in the same fashion as in the dictate study. First, consensus is expected to have little or no effect upon second-grade children, so they are not expected to respond to the inhibitory cause/good outcome situation as they are expected to in the dictate study. Here, as in all external-cause conditions, second-grade children

are expected to respond to the level of consequences. In this study, fifth-grade children are expected to respond more like second-graders than eighth-graders, and to attribute responsibility according to the severity of outcomes. Furthermore, inhibitory and facilitative consensus are expected to be important factors for eighth-grade children when judging both bad and good acts. For example, it is expected that the facilitative consensus/bad outcome condition will lead to relatively more discounting (i.e., less responsibility will be attributed, less diagnosticity attributed, and less allocated sanctioning) than when a command serves as the facilitative external cause. Thus, it is proposed that consensus is a more plausible external cause than are dictates for older children. Moreover, as was suggested earlier, no information regarding consensus is expected to lead to comparable levels of discounting as facilitative consensus when eighth-graders judge good outcomes, and no information should compare to inhibitory consensus when they judge bad outcomes.

Summary of Hypotheses for Experiment 1:

1. Valence of outcome.

1a. More responsibility and diagnosticity are attributed for bad outcomes than for good outcomes.

1b. This valence-of-outcome effect is particularly salient with second-grade children, compared to older children. Although older children also demonstrate such an effect, the

differences between valences are not as extreme as those demonstrated by younger children.

2. External cause.

2a. When children judge actors involved in bad outcomes, they attribute less responsibility and diagnosticity to them, and allocate them less punishment, when an external facilitative cause is present, than when the external cause is inhibitory or when no such information is available.

2b. This "discounting" effect is particularly salient in eighth-grade children, compared to younger children.

2c. When children judge actors involved in good outcomes, they attribute more responsibility and diagnosticity to them, and allocate them more reward, when an external inhibitory cause is present, than when the external cause is facilitative or when no such information is available.

2d. This "augmenting" effect is particularly salient in eighth-grade children, compared to younger children.

2e. When second-grade children are exposed to dictate information that is inhibitory to a good outcome, they demonstrate decreased attributions of responsibility, and allocate fewer rewards, than when no information regarding dictates is available.

2f. When second-grade children are exposed to dictate information that is facilitative to a good outcome, they demonstrate increased attributions of responsibility, and allocate more rewards, than when no information regarding dictates is available.

3. Severity of outcome.

3a. More responsibility and diagnosticity are attributed when outcomes are relatively severe rather than mild.

3b. This severity-of-outcome effect is particularly salient with second-graders, and decreases with older children.

Summary of Hypotheses for Experiment 2:

All hypothese from Experiment 1, except for 23 and 24, are relevant to Experiment 2, with the following addenda:

4a. Second-grade children will show less differentiation on all dependent measures according to external-cause information, compared to Experiment 1.

4b. Eighth-grade children will show more differentiation on all dependent measures according to external-cause information, compared to Experiment 1.

Experiment 1

Method

Subjects and Design. One hundred and seventeen children, 39 from each of three grades (second, fifth and eighth) were recruited from Edmonton Public Schools to participate in this experiment. The particular schools providing subjects were located in the same general area of the city, although the elementary schools where second and fifth grade children participated were much closer in proximity than was the junior high to either. All schools were located in middle-class neighbourhoods consisting of a heterogeneous mixture of semi-professionals and skilled laborers. Parental permission was received before children were allowed to participate.

Mean age of second grade children (n = 17 boys, 22 girls) was 7 years and 8 months. Mean age of fifth grade children (24 boys, 15 girls) was 10 years and 10 months. Mean age of eighth grade children(21 boys, 18 girls) was 13 years and 11 months.

From class lists, children were randomly assigned to one of three external-cause conditions (i.e., facilitative, inhibitory and control). Therefore, with two levels of valence (good and bad) and two levels of severity (high and low), the overall design consisted of a 3 (Grades) X 3 (Cause) X 2 (Valence) X 2 (Severity) factorial design, with repeated measures on the latter two variables.

Materials. Four different stories were constructed that varied valence and severity of outcome. It was considered important during construction of stimulus stories both to control for the severity of outcomes across grades and between valences. That is, an attempt was made to present stories whose outcomes would in themselves not be perceived as more or less severe by children of different grades. Moreover, the valences of such outcomes had to be approximately equal in severity; e.g., the good act of high severity had to be perceived as comparable in "goodness" as the high-severe bad act was in "badness". Without these precautions it would be difficult to interpret any significant effects involving severity and valence, or any interaction involving grades with either.

To effect such controls, a pilot study (see Appendix 1) was conducted. Thirty children of ages similar to the target sample judged the goodness and badness of 12 outcomes on 4-point scales. Analyses of these data allowed the selection of two good outcomes (one high and one low in severity) and two bad outcomes (also of high and low severity) for inclusion in stimulus stories. The selected outcomes were as follows:

Good outcome/ high severity: "Someone giving all their savings, which is a lot of money, to poor children."

Good outcome/ low severity: "Someone helping a child on the playground tie their shoes."

Bad outcome/ high severity: "Someone taking a bicycle that doesn't belong to them."

Bad outcome/ low severity: "Someone taking a pencil that doesn't belong to them."

These outcomes were combined with prefatory parental dictate information (which was either facilitative or inhibitory to the outcomes) for subjects receiving the external-cause manipulation. For control groups, the stories merely consisted of fictional children performing such acts, with no dictate information included. Thus, it was possible to construct three types of booklets containing stimulus stories. Each type of booklet contained all four outcomes (high - and low-severe bad and good acts), but differed from other types according to the particular external cause condition it represented. The general format of the booklets included the presentation of a story (e.g., "Steve's mother told Steve not to give any of his savings to poor children. But Steve gave all of his savings, which was a lot of money, to poor children."), followed by five dependent measures before the presentation of the next story. Each dependent measure occupied one page of the booklet, and consisted of a 4-point scale where the child could indicate his answer by marking a circle. A complete description of the booklets, including the exact wording of all stories and dependent measures, appears in Appendix 2.

Major dependent measures. The major dependent measures for Experiment 1 consisted of questions as to (1) how responsible each actor was for their particular outcome, (2) how often the actor performs similar actions (diagnosticity), and (3) how

much the actor should be rewarded (for good outcomes) or punished (for bad outcomes).

Manipulation checks. As a check on the effectiveness of the External Cause manipulation, children were asked how mad (in the case of bad outcomes) or how happy (for good outcomes) they thought the actors' mothers were when they found out about the story characters' behavior. It was reasoned that, if a mother told her child (e.g.) to steal a bicycle, then she should subsequently not be angry when the child actually did so.

Also, as a check on the severity manipulation, children were asked to make a moral judgment as to "how good" the actors were when they were involved in good outcomes, and "how bad" the bad-outcome actors were. While the data provided by this question were to serve primarily as a manipulation check, it was also expected to also provide some additional information regarding the moral judgment-responsibility attribution relationship.

Procedure. Subjects were tested in groups of two or three in a separate room located in their school. Instructions included telling the children that they would hear some stories about some boys their own age who go to another school, and afterward would be asked some questions about them. It was explained that one of the questions would be how responsible the boys were for the things that happen in the stories. At this point, the experimenter took several minutes to ensure that the subjects understood the concept of responsibility.

Specifically, the experimenter explained that "if someone is responsible for a good or a bad thing happening, we would say that it's his fault that the good or bad thing happened. So you see, it can be someone's fault that a good thing happened, the same way it can be their fault that a bad thing happened. If someone is responsible for a bad thing that happened, we might blame them for it. But if they're responsible for a good thing that happened, we might want to thank them instead."[2] Originally, an attempt was made to present the instructions with a tape recorder, but it soon became apparent that some of the younger children were not sufficiently attending to this crucial stage of the procedure. Therefore, the experimenter presented the instructions personally with each session. However, an attempt was made to minimize experimenter effects by giving the exact same instructions and definitions to all subjects in all grades. Fortunately, it was possible to present the subsequent stories and experimental manipulations with the tape recorder, as the children were usually sufficiently interested in the experiment after presentation of the instructions to attend to the material.

Following the introduction and the discussion of responsibility, subjects were shown a sample booklet and told how to use the scales. Specifically, it was made clear that (e.g.) "very responsible" (4) was more responsible than "pretty responsible" (3), which in turn was more responsible than "a little bit responsible" (2). Even the youngest children demon-

strated no apparent difficulty understanding these distinctions.

The children were then told that they would hear the stories on the tape recorder and that they could read the stories on their booklets while they were being spoken. Booklets were then distributed (subjects were, of course, in the same treatment condition). The experimenter than played the first story (each story was told twice on the tape), ensured that the children understood the story, then proceeded to play the first question. The time interval between questions was held constant at seven seconds. Each experimental session lasted approximately 15 - 20 minutes.

Upon completion of a session, subjects were told that the experimenter would come to their classroom after the experiment was over, to explain more about it and to answer their questions. It was emphasized that it was very important that they not discuss the experiment with their classmates until everyone had participated.

Results

Data analysis. A total of 129 subjects participated in the experiment, but data for 12 subjects were randomly excluded in order to equalize the number of subjects per cell at 13.

While it is unusual for sex to operate as a significant factor in moral judgment research (cf. Wright, 1971),[3] a preliminary Sex X Grades X Valence analysis of variance on responsibility scores was performed to ensure that the subsequent analyses,

to be collapsed over the sex variable, would not obscure possibly significant effects for sex. This unequal-n, unweighted means analysis revealed no main effect for sex or interaction with grades and/or valence (F's < 1). While this analysis, of course, pertains only to attributions of responsibility, it should be noted that sex differences in cell means on all dependent variables rarely differed more than small amount, and never in a consistant fashion. Consequently, in order to benefit by additional statistical power, the sex variable was collapsed for all subsequent analyses.

External-cause manipulation check A potential problem with the credibility of the external-cause manipulation had been anticipated. Specifically, it could be argued that any failure to find predicted or unpredicted significant effects according to parental dictate information may be due to subjects' failure to believe such information, particularly dictates which are facilitative of bad outcomes and inhibitory of good ones. Therefore, the post-experimental interview for some randomly chosen subjents in these latter conditions included informal questions as to possible explanations the children might have for such parental behavior. Subjects of all grades were able to suggest explanations, although these explanations themselves varied widely in credibility. Generally, however, the legitimacy of the facilitative/ bad outcome dictate was rationalized (e.g., "The mother was poor and the boy needed the bicycle to get to school"), while ulterior and apparently acceptable motivations

were usually attributed to mothers who gave the inhibitory/good outcome dictate (e.g., "She wanted him to save his money," rather than give it away). Although these informal answers typically indicated that subjects perceived the dictate information as credible, it should be added that the more objective data for attributions of anger and happiness to mothers, to be described shortly, also address the issue of the credibility of the external-cause manipulation.

Separate analyses of variance were performed on "how mad" and "how happy" the mothers were estimated to be. For the first category, a highly significant main effect for external cause, F (2,108) = 39.76, $p < .01$, indicated that this information had produced the intended effect for the bad-outcome stories. Subjects in the facilitative-cause condition attributed the least anger to the mother (M = 2.0), while subjects in the inhibitory-cause condition attributed the most (M = 3.38). Overall, the control condition resulted in scores which did not differ significantly (M = 3.18) from the inhibitory-cause condition, according to a Duncan's test of means, but this result will be qualified in the discussion of the Cause X Severity interaction.

On the "how mad" scores, there was also a highly significant main effect for severity of outcomes, F (1,108) = 120.7, $p < .01$. As would be expected, mothers were judged to be more angry when outcomes were relatively severe (i.e., the stolen bicycle; M = 3.27) rather than mild (i.e., the stolen pencil; M = 2.44). This finding also attests to the effectiveness of

the severity manipulation.

As mentioned previously, there was also a significant Cause X Severity interaction, F (2,108) = 17.45, $p < .01$. The cell means for this interaction (see Table 1) indicated that the obtained pattern of results for the cause main effect held particularly for the high-severity condition. That is, subjects in the high-severity condition estimated almost identical levels of anger in mothers when the dictate had been inhibitory and when no dictate information was given, with less anger attributed in the facilitative-cause condition. In contrast, subjects in the low-severity condition did estimate higher anger when the parental dictate had been inhibitory, compared to no-information (control) subjects. These differences were significant at the .05 level according to Duncan's multiple range test on cell means.

TABLE I

Mean Estimated Anger of Actor's Mother According to Severity of Bad Act and External Cause (Dictate)

	External Cause		
Severity	Facilitative	Inhibitory	Control
High	2.13_a	3.82_b	3.85_b
Low	1.87_c	2.95_d	2.51_e

Note. Higher means indicate higher attributed anger. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

There was also a significant main effect for grades, F (2,108) = 6.8, $p < .01$. Second-graders attributed the most anger to mothers (M = 3.18) overall, while eighth-graders attributed the least (M = 2.56). The scores for fifth-graders (M = 2.82) were not significantly higher than those of eighth-graders, but were significantly lower than those of second-grade subjects ($p < .01$).

Grades did not interact significantly with external cause; therefore there is no indication that the cause manipulation had any differential effect among grades. The complete table for the analysis of variance on the "how mad" scores appears in Appendix 3.

For the "how happy" scores, there was a highly significant main effect for external cause, F (2,108) = 173.6, $p < .01$, which indicates that the cause manipulation had been salient to the children. Subjects in the facilitative-cause condition estimated the mothers to be happiest in response to the good act (M = 3.62), while subjects in the inhibitory-cause condition estimated the mothers to be least happy (M = 1.36). Subjects in the control group estimated a relatively high degree of happiness in the mothers (M = 3.22), but significantly less ($p < .05$ by the Duncan test) than facilitative-cause subjects.

While there was a moderately significant main effect for severity, F (1,108) = 6.31, $p < .05$, with more happiness attributed to mothers in the high- than in the low-severe condition (Ms = 2.83 and 2.63, respectively), there was also a Severity

X Cause interaction, F (2,108) = 8.39, $p < .01$, which indicated that the severity effect was limited entirely to the no-information (control) subjects (see Table 2).

The main effect for grades did not reach an acceptable level of significance, nor were any other interactions significant. The summary of the analysis of "how happy" scores appears in Appendix 4.

Judgments of Bad. Analysis of variance (see Appendix 5) revealed that judgments of "bad" regarding story characters were affected by severity, F (1,108) = 42.0, $p < .01$, and external cause, F (2,108) = 4.58, $p < .05$. The child who committed the high-severe act (stealing a bicycle) received a higher "badness" score (M = 3.69) than did the child in the low-severity story (M = 3.18). The significant external cause main effect was due to decreased judgments of badness in the facilitative-cause condition (M = 3.23), compared to control subjects (M = 3.49) and increased badness scores in the inhibitory-cause condition (M = 3.59).

However, these variables (severity and external cause) also yielded a significant interaction, F (2,108) = 3.92, $p < .05$. Cell means for this interaction (see Table 3) indicated that the pattern of results for the external-cause main effect was restricted mainly to the high-severity condition, with little differentiation by cause in the low-severity condition. That is, only subjects in the high-severity condition judged the actor to

TABLE 2

Mean Estimated Happiness of Actor's Mother According to Severity of Good Act and External Cause (Dictate)

	External Cause		
Severity	Facilitative	Inhibitory	Control
High	3.64_a	1.31_b	3.54_a
Low	3.59_a	1.41_b	2.90_c

Note. Higher means indicate more attributed happiness. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

TABLE 3

Mean Judgments of "Bad" According to External Cause (Dictate) and Severity of Outcomes

	External Cause		
Severity	Facilitative	Inhibitory	Control
High	3.33_a	3.95_b	3.79_b
Low	3.13_c	3.23_{ac}	3.18_{ac}

Note. Higher means indicate increased judgments of "bad". Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

be less bad when the dictate had been facilitative than when it had been inhibitory or when no dictate information was given. Subjects in the low-severity condition showed no significant differentiation by external cause. Differences were significant at the .05 level by the Duncan test.

Also, there was a significant Grades X Cause interaction, F (4,108) = 2.69, $p < .05$. Between-grades comparisons on cell means (see Table 4) indicated that there were significant differences between grades in the facilitative-cause condition, but not in the inhibitory-cause and control conditions. Therefore, when the parental dictate was facilitative to the bad outcome, second-grade children judged the actor more harshly than did fifth-graders, who in turn judged the actor more harshly than did eighth-grade children (ps $< .05$ by the Duncan test). Within-grades comparisons indicated that fifth- and eighth-grade subjects, but not second-graders, judged actors to be also less bad in the facilitative-cause condition compared to the respective control groups. There was no evidence that external-cause information had any effect on second-graders' judgments of badness.

No other main effects or interactions reached significance.

Judgments of Good. The analysis of variance on the children's judgments of "good" regarding good-outcome actors produced significant differences for all main effects and most interactions. First, there was a main effect for severity, F (1,108) = 10.97, $p < .01$. The actor in the high-severity condition

TABLE 4

Mean Judgments of "Bad" According to Grades and External Cause (Dictate)

	External Cause		
Grades	Facilitative	Inhibitory	Control
2	3.62_{a}	3.69_{a}	3.38_{ab}
5	3.15_{bc}	3.62_{a}	3.42_{ab}
8	2.92_{c}	3.46_{ab}	3.65_{a}

Note. Higher means indicate increased judgments of "bad". Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

received higher "goodness" scores overall (M = 3.61) than did actors in the low-severity story (M = 3.37). The main effect for cause, F (2,108) = 5.01, $p < .01$, indicated that actors in inhibitory-cause conditions were not judged quite as good (M = 3.27) as actors in facilitative-cause (M = 3.67) and control (M = 3.53) conditions. However, this main effect in particular will be qualified by the discussion of the Grades X Cause interaction.

Severity and cause also interacted, F (2,108) = 4.38, $p < .05$. This interaction indicated that the severity main effect was limited largely to the control-group subjects. However,

further statistical exploration of this interaction was deferred, as severity and cause were also involved in a higher-order interaction with grades.

The main effect for grades, F (2,108) = 4.63, $p < .05$, is qualified by a significant Grades X Cause interaction, F (4,108) = 7.54, $p < .01$. Cell means for this interaction (see Table 5) indicated that second-grade subjects judged actors in the inhibitory-cause condition to be less "good" than did fifth- and eighth-grade subjects, thus supporting predictions. Also, it had been predicted that eighth-graders in the facilitative-cause

TABLE 5

Mean Judgments of "Good" According to Grades and External Cause (Dictate)

	External Cause		
Grade	Facilitative	Inhibitory	Control
2	3.81_a	2.54_b	3.50_a
5	3.85_a	3.62_a	3.54_a
8	3.35_a	3.65_a	3.54_a

Note. Higher means indicate increased judgments of "good". Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

condition would attribute less goodness to actors than would

second- and fifth-grade subjects, due to the discounting of responsibility by eighth-graders. The Duncan test was marginally significant ($p < .06$) for both between-grades comparisons, thus lending some qualified support for this prediction.

A significant Grades X Severity X Cause interaction, F (4,108) = 2.47, $p < .05$, is difficult to interpret (see Figure 1). However, this higher-order interaction does not seem to substantially qualify the Grades X Cause interaction.

Finally, it should be noted that these results for both badness and goodness judgments indicate that the severity manipulation had the intended effect. More importantly, however, control group data indicated that severity in both valence conditions was judged comparably among grades. Therefore, analyses of the major dependent variables were performed without prejudice.

Attributions of Responsibility. Analysis of variance revealed a significant main effect for grades, F (2,108) = 8.09, $p < .01$, on attributions of responsibility. Fifth-grade subjects attributed the most responsibility overall (M = 3.54), while second- and eighth-grade subjects attributed comparable levels of responsibility overall (Ms = 3.11 and 3.12, respectively). There was also a highly significant main effect for external cause, F (2,108) = 18.46, $p < .01$. Most responsibility was attributed when the nature of the parental dictate was inhibitory to the outcome (M = 3.62), and least whem facilitative (M = 2.88). Children receiving no information about parental dictates attributed responsibility amost exactly intermediately

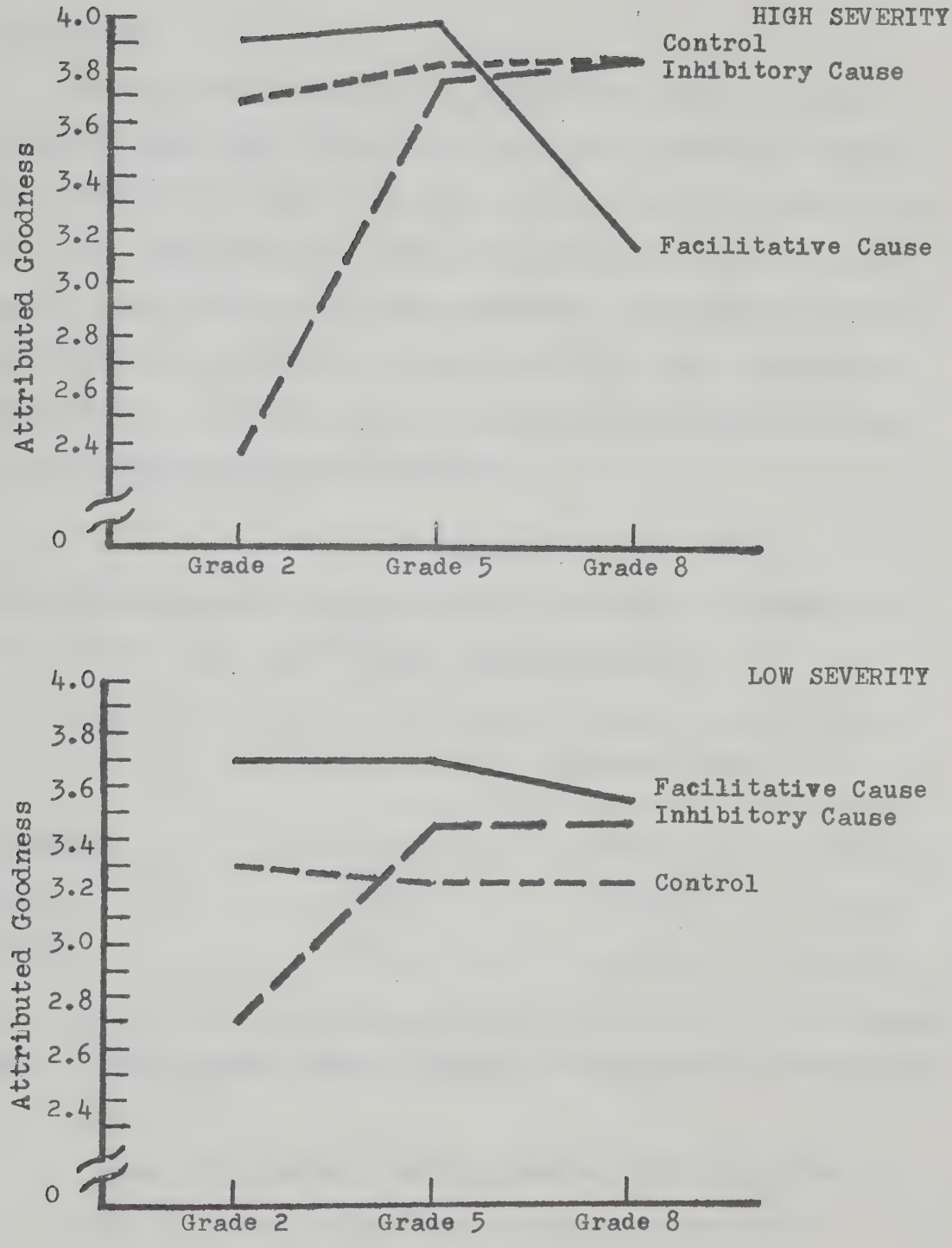

Figure 1. Mean Judgments of "Good" According to Grades, Severity of Outcome, and External Cause (Dictate).

(M = 3.26).

However, despite the highly significant nature of the external cause main effect, this result is qualified somewhat by a significant Cause X Valence interaction, F (2,108) = 5.14, $p < .01$. Cell means for this interaction (see Table 6) indicated that, in the bad-outcome condition, children treated no information as comparable to an inhibitory cause, whereas no information was comparable to the facilitative-cause information in the good-outcome condition.

TABLE 6

Mean Attributions of Responsibility According to Valence of Act and External Cause (Dictate)

	External Cause		
Valence	Facilitative	Inhibitory	Control
Bad	2.73_a	3.72_b	3.53_b
Good	3.03_a	3.52_b	3.00_a

Note. Higher means indicate higher attributions of responsibility.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

On the other hand, it had been predicted that grades would interact with valence and cause. While this interaction (see

Table 7) was not statistically significant, specific predictions regarding this interaction were supported by the Duncan test. In particular, this analysis revealed that the pattern of the Valence X Cause interaction was demonstrated most saliently by eighth-grade subjects. That is, the oldest children attributed the least responsibility for bad outcomes when there was a facilitative cause present (compared to an inhibitory cause and no information), and the most responsibility for good outcomes in the inhibitory-cause condition (with no information and facilitative-cause information treated approximately the same). To facilitate comprehension, mean values for these comparisons are plotted in Figure 2.

It had also been predicted that eighth-grade subjects would discount responsibility for good outcomes when the external cause was facilitative. While the Duncan test did not reach significance, the less conservative t-test for multiple comparisons (considered appropriate for a priori predictions) did indicate a significant difference between eighth-grade subjects in the facilitative-cause/good-outcome cell and the combined means of younger subjects in the same condition, t (108) = 2.05, $p < .05$, one-tailed.

External cause also interacted with the severity of outcomes, F (2,108) = 7.63, $p < .01$. Cell means for this interaction (see Table 8) indicated that most responsibility was attributed in the high-severity/inhibitory-cause condition, and least in the

TABLE 7

Mean Attributions of Responsibility According to Grades, Valence, and External Cause (Dictate)

Grade	External Cause		
	Facilitative	Inhibitory	Control
	Bad Outcome		
2	2.35_{a}	3.62_{efg}	3.54_{defg}
5	3.31_{cdefg}	3.92_{g}	3.73_{fg}
8	2.54_{ab}	3.62_{efg}	3.31_{cdefg}
	Good Outcome		
2	3.15_{bcdef}	3.15_{bcdef}	2.85_{abcd}
5	3.27_{bcdefg}	3.76_{fg}	3.23_{bcdefg}
8	2.65_{abc}	3.65_{efg}	2.92_{abcde}

Note. Higher means indicate higher attributions of responsibility.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

high-severity/facilitative-cause condition. Interestingly, more responsibility was attributed in the facilitative-cause condition when severity was low rather than high, whereas the remain-

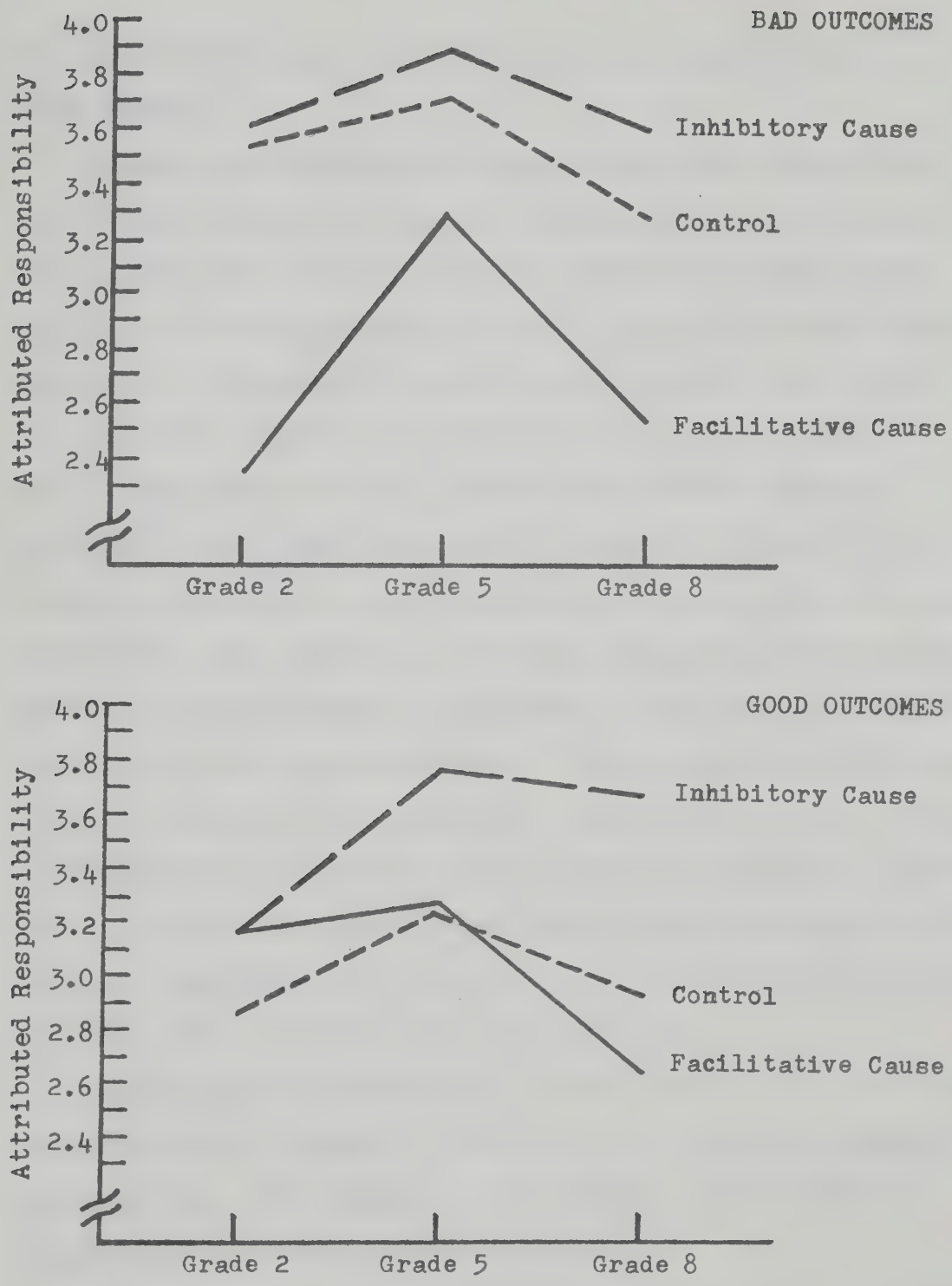

Figure 2. Mean Attributions of Responsibility According to Grades, Valence, and External Cause (Dictates).

ing cause conditions produced the predicted severity-of-outcome effect.

However, the significant Grades X Severity X Cause interaction, F (4,108) = 2.57, $p < .05$, indicated that this latter interaction was evidenced solely by fifth- and eighth-grade children, as second-graders did not differentiate their attributions of responsibility according to external cause in the low-severity condition (see Figure 3). The Duncan test on these means indicated that, in the low-severity condition, children in all grades apparently treated no information as to cause as comparable to facilitative-cause information, whereas only fifth- and eighth-grade children (and not second-graders) greatly increased their attributions of responsibility when the parental dictate was inhibitory to the low-severity outcome (ps $< .01$), compared to the respective control-group means. Also, within-grades comparisons revealed that only second- and eighth-grade children (and not fifth-graders) showed increased attributions of responsibility in the facilitative-cause condition when outcomes were low rather than high (ps $< .05$).

The analysis of variance for the main effects of valence and severity did not approach significance, nor did the remaining interactions. The summary of the analysis of attribution of responsibility scores appears in Appendix 7.

Diagnosticity. Analysis of variance on subjects' estimates of how often story characters demonstrated similar behaviors revealed a significant main effect for valence, F (1,108) = 7.59,

TABLE 8

Mean Attributions of Responsibility According to Severity of Outcome and External Cause (Dictate)

	External Cause		
Severity	Facilitative	Inhibitory	Control
High	2.73_a	3.74_b	3.40_c
Low	3.03_d	3.50_c	3.13_d

Note. Higher means indicate higher attributions of responsibility.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

$p < .01$. Actors performing good acts were generally seen as more likely to repeat such behavior (M = 2.60) than were actors who did bad acts (M = 2.38). However, a significant Valence X Severity interaction, F (1,108) = 5.03, $p < .05$, indicated that the effect for valence was particularly salient when severity was low rather than high (see Table 9).

It had been predicted that grades would interact with external cause, with older children attributing more diagnosticity with an inhibitory cause. While the overall interaction itself did not reach statistical significance, a Duncan test on cell means did lend partial support for this expectation (see Table

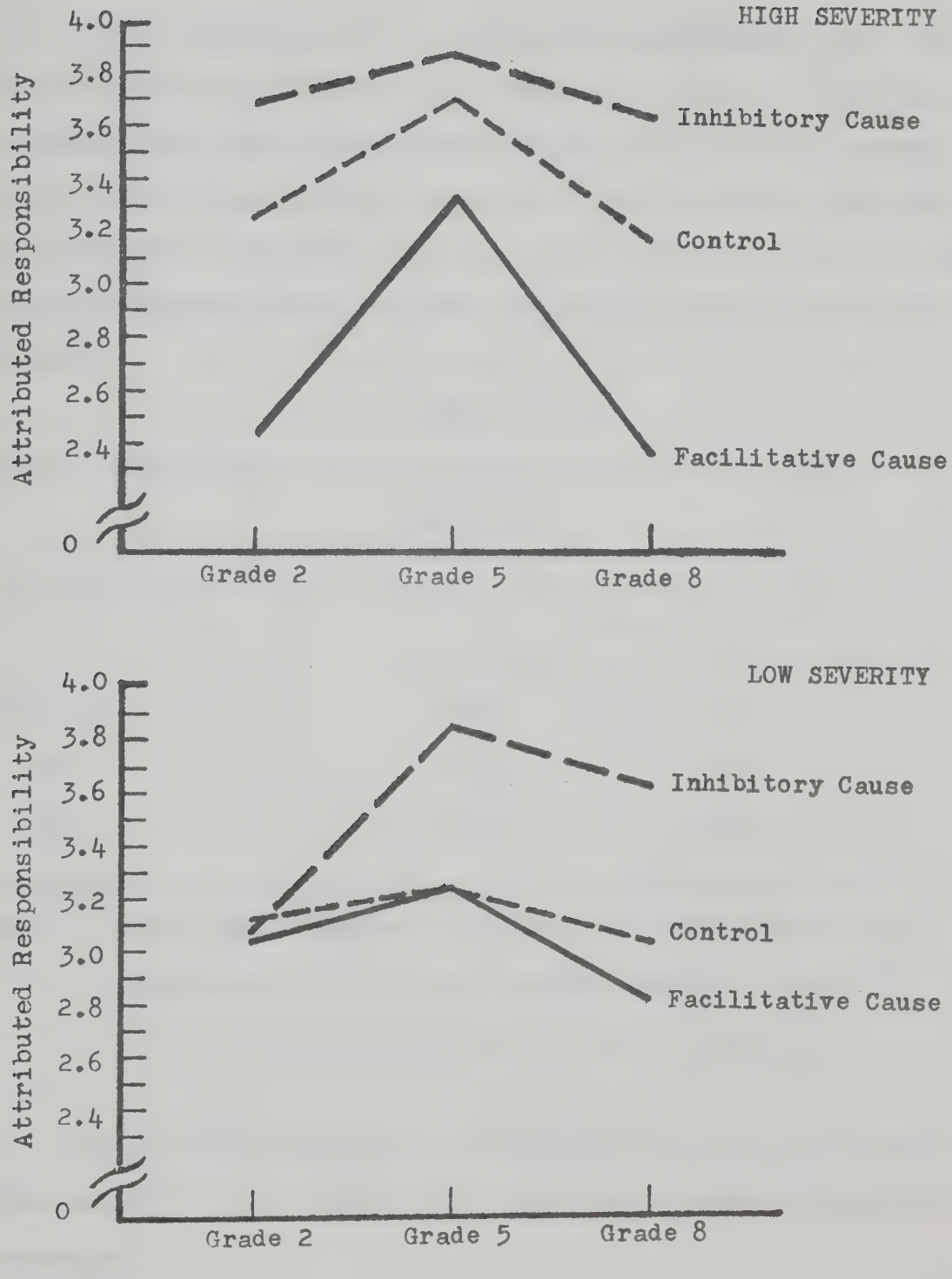

Figure 3. Mean Attributions of Responsibility According to Grades, Severity, and External Cause (Dictate).

10). While eighth-grade subjects in the inhibitory-cause condition did not show increased diagnosticity compared to controls, eighth-grade facilitative-cause subjects did attribute significantly less diagnosticity, compared to eighth-graders receiving no information regarding external cause. Children in the second- and fifth-grades did not demonstrate this apparent discounting effect.

TABLE 9

Mean Diagnosticity Scores According to Valence and Severity of Outcomes

	Severity	
Valence	High	Low
Bad	2.44_{ab}	2.32_{a}
Good	2.55_{bc}	2.65_{c}

Note. Higher means indicate higher attributed diagnosticity. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

No other main effects or interactions reached significance. The summary of the analysis of diagnosticity scores appears in Appendix 8.

Allocated Punishment. As in the judgments of badness, this analysis revealed a highly significant main effect for severity, F (1,108) = 141.3, $p < .01$. As would be expected, the high-

severe story character was allocated more punishment (M = 3.5) than was the low-severe character (M = 2.52). Unlike the analysis of "bad" scores, there was also a main effect for grades, F (2,108) = 5.14, $p < .01$. Second-grade subjects allocated significantly more punishment overall (M = 3.27) than did fifth-grade subjects (M = 2.94) and eighth-grade subjects (M = 2.82; $ps < .05$ by the Duncan test). The difference between the latter two grades did not approach significance.

TABLE 10

Mean Diagnosticity Scores According to Grades and External Cause (Dictate)

	External Cause		
Grade	Facilitative	Inhibitory	Control
2	2.67_{ab}	2.21_{a}	2.52_{ab}
5	2.58_{ab}	2.46_{ab}	2.63_{ab}
8	2.23_{a}	2.41_{ab}	2.71_{b}

Note. Higher means indicate higher attributed diagnosticity. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

There was also a significant main effect for external cause, F (2,108) = 3.63, $p < .05$. While least punishment was allocated overall in the facilitative-cause condition, this main effect is

qualified by a significant Grades X Cause interaction, F (4,108) = 3.22, $p < .05$. Cell means for this interaction (See Table 11) indicated that it was mainly the older children, particularly eighth-graders, who showed decreased allocated punishment in the facilitative-cause condition. As predicted, second-grade subjects did not seem to be affected by external-cause information when allocating punishments for bad acts.

The summary of the analysis of variance on punishment scores appears in Appendix 9.

TABLE 11

Mean Allocated Punishment According to Grades and External Cause (Dictate)

	External Cause		
Grade	Facilitative	Inhibitory	Control
2	3.35_{a}	3.00_{ab}	3.46_{a}
5	2.77_{b}	3.12_{ab}	2.92_{ab}
8	2.27_{c}	3.04_{ab}	3.15_{ab}

Note. Higher means indicate higher levels of allocated punishment.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

Allocated Reward. Analysis of variance on the scores repres-

enting the degree to which children perceived the good-outcome actors to be deserving of rewards (see Appendix 10) revealed a highly significant main effect for outcome severity, F (1,108) = 41.47, $p < .01$. As expected, the high-severe actor (who gave his savings to poor children) was rated higher in reward-deservingness (M = 2.87) than was the low-severe actor (who helped a child with his shoelaces; M = 2.45). There was also a significant main effect for cause, F (2,108) = 4.86, $p < .01$, with subjects receiving facilitative-cause information rating actors generally more deserving of rewards (M = 2.91) than did subjects who received inhibitory-cause information (M = 2.31) or no information (M = 2.46).

TABLE 12

Mean Allocated Reward According to Grades and External Cause (Dictate)

	External Cause		
Grade	Facilitative	Inhibitory	Control
2	3.27_{a}	2.19_{cd}	2.92_{abc}
5	3.08_{ab}	2.12_{d}	2.50_{bcd}
8	2.38_{bcd}	2.62_{abcd}	1.96_{d}

Note. Higher means indicate higher levels of allocated reward. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

However, the main effect for cause is qualified by a significant Grades X Cause interaction, F (2,108) = 2.91, $p < .05$. As Table 12 indicates, it was mainly the younger children, especially second-graders, who showed the increased allocated rewards in the facilitative-cause condition. As predicted, eighth-grade subjects perceived significantly less reward-deservingness in actors who performed good acts after being told to do so (i.e., facilitative-cause information) than did second-graders. This result coincides with the finding that eighth-graders in the facilitative-cause condition had also judged actors to be less "good" than did the younger subjects, as discussed earlier. On the other hand, it should be noted that this obtained pattern of results for allocated reward also occured in the Control (i.e., no-information) condition. Thus, while eighth-grade control subjects had scored comparably to second-grade controls on judgments of good, they nevertheless perceived the actors as less deserving of rewards than did the second-graders.

The means for the simple effect for grades in the inhibitory-cause condition were in the predicted directions, with eighth-grade subjects scoring highest, but differences did not reach acceptable levels of significance by either the Duncan or the t-test.

When the within-grades effect for cause is considered, the Duncan test ($p < .06$) and the t-test (t (108) = 2.09, $p < .05$, one-tailed) indicate a tendency for second-grade subjects to allocate less reward in the inhibitory-cause condition than in the control condition. This finding also coincides with the

obtained pattern of results of "good" scores and supports the hypothesis that children of this age have a concept of good behavior that is oriented toward obedience.

Similarly, in contrast to the results of second-grade subjects, eighth-graders allocated more reward to actors in the inhibitory-cause condition than did control subjects, t (108) = 1.89, $p < .05$, one-tailed. This result was also predicted.

All other main effects and interaction in this analysis failed to reach significance.

Discussion

The results of Experiment 1 provide some support for most of the hypotheses. First, it was found that children not receiving external-cause manipulations generally attributed more responsibility to actors involved in bad outcomes than in good outcomes, and further that the valence-of-outcome effect was significantly more salient with the younger children than with older children. The proposed explanation for this phenomenon, i.e., that facilitative causes are normally assumed in the occurrence of good outcomes, and inhibitory causes are assumed in the case of bad outcomes, was supported by the significant Valence X Cause interaction on attributions of responsibility.

However, the diagnosticity results did not support the valence-of-outcome hypotheses. In fact, children of all grades rated the good-outcome actors as more likely to repeat similar

behaviors than the bad-outcome actors. This finding is interesting because, as noted previously, actors were generally held less responsible for good outcomes than for bad outcomes, at least by control-group subjects. It seems likely that children were invoking a priori beliefs that people are generally more inclined to perform good acts than bad acts, thus overriding any tendency for consistency between attributions of responsibility and diagnosticity.

Again, while it was apparent that when children were judging good acts, no information as to external cause was comparable to facilitative-cause information, there was a slight but significant tendency for eighth-grade subjects to discount still more responsibility than younger subjects whem the external cause (parental dictate) was facilitative to a good outcome. This finding is consistent with predictions. However, contrary to predictions, when outcomes were bad, eighth-grade subjects did not seem to discount responsibility more than second-graders when the cause was facilitative. Both grades showed a marked decrease in attributed responsibility in the facilitative-cause/bad-outcome condition, compared to their respective control groups. Curiously, it was the fifth-grade subjects--not the second-graders--who attributed most responsibility (although still less than did comtrol-group subjects) in this condition.

In addition, the expected age differences were strongly supported by data on the remaining dependent measures. First, eighth-graders attributed less diagnosticity when the parental

dictate was facilitative, compared to controls, regardless of valence of outcome. Second- and fifth-grade subjects did not demonstrate a comparable decrease in diagnosticity in the facilitative-cause condition. More interesting, perhaps, are scores for moral judgments and sanctioning. When outcomes were bad, there was a strong developmental trend toward decreased judgments of bad and allocated punishment with increasing age, in the facilitative-cause condition. The finding that second-graders in this condition judged actors more harshly than did older children, despite the fact that such actors were behaving in accordance with parental dictates, provides a replication of results by Berg-Cross (1975). However, why second-graders in the present study nevertheless attributed less responsibility to actors in the facilitative-cause/bad-outcome condition is not known.

Responses by second-grade children in the facilitative-cause/<u>good</u>-outcome condition provide support for the hypothesis that the concept of good held by young children is based on obedience, rather than being oriemted toward a recipient's feelings or welfare. Second-grade children actually judged actors to be more "good" and allocated them more reward when they performed good acts in response to a facilitative parental dictate than when they were given no such information. Curiously, this increase occured despite a non-significant increase in attributions of responsibility by these second-graders. Again, there is little recognizable consistency between second-graders' attributions and their moral judgments.

Considering now the observed effects of an inhibitory external cause on attributions of responsibility, it is clear that the results support the predicted fashion in which second-grade subjects would respond to a good outcome that is counter to an inhibitory parental dictate. First, only fifth- and eighth-grade subjects showed a significant augmenting of responsibility in the inhibitory-cause/good-outcome condition, compared to controls. Moreover, these same fifth- and eighth-grade children also judged actors to be more "good" and allocated them more reward as well. Second-grade children actually judged actors in the inhibitory-cause/good-outcome condition to be less good and less deserving of reward than did control-group second-graders. These latter findings coincide with those of second-graders in the facilitative-cause/good-outcome condition, discussed previously, in their support for the hypothesized obedience-oriented concept of good held by young children.

The main unsupported prediction regarding external cause was that age would interact with inhibitory cause on judgments of diagnosticity. Although eighth-graders did perceive less diagnosticity when the parental dictate was facilitative to the bad or good outcome, they did not show the predicted increase in diagnosticity whem the dictate was inhibitory.

Finally, hypotheses pertaining to the severity of outcomes received only qualified support. The Grades X Severity X Cause interaction on attributions of responsibility indicated that the predicted severity-of-outcome effect occurred in an incon-

sistent fashion which is difficult to interpret. Overall, second graders were most affected by the severity manipulation, as their attributions of responsibility were practically undifferentiated by external cause when severity was low rather than high.

The one consistent (although unpredicted) finding was that children in all grades showed increased attributions of responsibility when the external cause was facilitative to the low-severity rather than high-severity outcome. It is possible that these children may have perceived less latitude of choice in actors ordered by parents to do high- rather than low-severe acts, thereby leading to relatively more responsibility attributed in the latter condition.

The only other significant finding pertaining to outcome severity was the Severity X Valence interaction on judgments of diagnosticity. This interaction indicated that the low-severe/good act was given higher diagnostic value than both the low-severe/bad act and the high-severe/bad act. Moreover, the actor performing the low-severe good act (i.e., helping a child with his shoelaces) was perceived as marginally more likely to perform other good acts than was the actor performing the high-severe good act. This finding coincides with the result, discussed above, that a facilitative cause was apparently assumed in the occurrence of good outcomes, and the further finding that more responsibility was attributed when the cause was facilitative to a low- rather than a high-severe act. It is tempting

to generalize from this pattern of results and suggest that they support the intuitively plausible hypothesis that good deeds of low severity appear more spontaneous and are less likely to arouse suspicion as to ulterior motives than are good deeds of a grander scope. Conversely, bad deeds of lessor consequence may more readily be attributed to a temporary lapse of moral vigilance than bad acts of higher magnitude. This interaction between the valence and the severity of acts would seem to warrant consideration in future research.

Experiment 2

Method

Subjects and design. One hundred and twenty-six children, 42 from each of three grades (second, fifth and eighth), participated in Experiment 2. Eighth-grade subjects were drawn from the same junior high as that providing subjects for Experiment 1, whereas different elementary schools provided second- and fifth-grade subjects. However, these elementary schools were in the same general vicinity as those schools used in the first Experiment. Parental permission was received before children were allowed to participate.

Mean age of second-grade children was 7 years and 7 months. Mean age of fifth-grade children was 10 years and 11 months. Mean age of eighth-grade children was 13 years and 11 months. Ages of subjects in each grade for Experiment 2 were therefore comparable to the ages of subjects in Experiment 1. Sex of subjects was evenly divided in all grades.

The method for assigning subjects to conditions and the overall design were identical to those of Experiment 1.

Materials and procedure. The same four stories from Experiment 1 were used in this experiment, with one important variation, i.e., the nature of the external cause information. Facilitative or inhibitory peer consensus information, rather than parental dictates, served as the external cause manipulation for Experiment 2. For example, the high-severe bad act performed in the presence of facilitative consensus read:

"Almost all of the kids in Bobby's school have sometimes taken bicycles that didn't belong to them, and kept them for themselves. Bobby took another boy's bicycle and kept it for himself." A complete description of the stories appears in Appendix 11.

The same major dependent measures from Experiment 1 (responsibility, diagnosticity, sanctioning) were included and were worded identically. However, the manipulation check for Experiment 2 consisted of a question as to how many other children in the respective actors' schools would have performed a similar act, "if they had the chance." Also, because the analysis of the moral judgment data in Experiment 1 had demonstrated that the severity manipulation had had its intended effect, it was not considered necessary to repeat the "how bad" and "how good" questions in Experiment 2. Therefore, these questions were deleted from the booklets. With these few exceptions, the booklets were identical to those used in Experiment 1.

Finally, the procedure was identical to that followed in Experiment 1.

Results

Data analysis. One hundred and twenty-eight subjects participated in Experiment 2. However, data for two second-grade subjects had to be discarded for inability by the experimenter to make instructions clear. The final analysis therefore yielded a 3 (Grades) X 3 (Cause) X 2 (Valence) X 2 (Severity)

factorial design, with repeated measures on the last two variables. However, as Valence was not a factor in the analyses of scores for allocated punishment and reward, separate Grades X Cause X Severity analyses were performed on these data.

Manipulation checks. Analysis of variance on the manipulation-check data (see Appendix 12) indicated that the consensus manipulation had been highly salient to the children, F (2,117) = 66.79, $p < .01$. Subjects in the facilitative-cause (high consensus) condition estimated the highest proportion of other children who would act similarly (M = 2.97), while subjects in the inhibitory-cause (low consensus) condition estimated the lowest (M = 1.58). Control (no-information) subjects scored intermediately (M = 2.06).

There was also a main effect for severity, F (1,117) = 9.31, $p < .01$, which indicated that more children were estimated to perform low-severe acts (M = 2.28) than high-severe acts (M = 2.13). Severity interacted significantly with external cause, F (2,117) = 5.74, $p < .01$, which indicated that the main effect for severity was restricted mainly to control-group subjects, as subjects receiving consensus information did not differentiate their estimates by severity. However, severity also interacted significantly with grades, F (2,117) = 14.44, $p < .01$. Cell means for this interaction (Table 13) indicated that eighth-grade subjects made lower estimates than did second-graders when severity was high, and higher estimates than did second-graders when severity was low. These differences are significant at

the .05 level by the Duncan test.

A significant main effect for valence, $\underline{F}$ (1,117) = 6.73, $\underline{p} < .05$, indicated that subjects generally estimated higher consensus overall for good acts (M = 2.32) than for bad acts (M = 2.10). But a highly significant Grades X Valence interaction, $\underline{F}$ (2,117) = 23.66, $\underline{p} < .01$, indicated that this "Pollyanna" effect was demonstrated almost solely by second-grade subjects. In fact, eighth-grade subjects responded almost exactly opposite to second-graders, and estimated that proportionately more children would engage in bad acts than good acts ($\underline{p} < .05$ by the Duncan test).

There was also a significant Valence X Severity interaction, $\underline{F}$ (1,117) = 10.08, $\underline{p} < .01$. Further analysis of this interaction revealed that the high-severe/good act (M = 2.32) was given higher consensus estimates than the high-severe/bad act (M = 1.94; $\underline{p} < .05$), but that there was no simple effect for valence in the low-severity condition (Ms = 2.25 and 2.34 for bad and good acts, respectively).

Neither the main effect for grades nor the Grades X Cause interaction approached significance.

Attributions of Responsibility. Analysis of variance (see Appendix 13) revealed a significant main effect for grades, $\underline{F}$ (2,117) = 3.79, $\underline{p} < .05$. As in Experiment 1, fifth-grade subjects attributed the most responsibility overall (M = 3.60), while eighth-grade subjects attributed the least responsibility overall (M = 3.23). The scores for second-grade subjects (M =

3.40) were intermediate.

TABLE 13

Mean Estimates of Consensus According to Grades and Severity of Outcomes

	Severity	
Grade	High	Low
2	2.33_{ab}	2.14_{b}
5	2.14_{b}	2.32_{ab}
8	1.90_{c}	2.38_{a}

Note. Higher means indicate higher estimates of consensus. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

Unlike Experiment 1, there were significant main effects for valence, $\underline{F}$ (1,117) = 10.58, $\underline{p} < .01$, and severity, $\underline{F}$ (1,117) = 11.56, $\underline{p} < .01$. More responsibility was attributed to actors involved in bad acts (M = 3.56) than in good acts (M = 3.27), and more responsibility was attributed to actors involved in high-severe acts (M = 3.52) than low-severe acts (M = 3.31).

The Valence X Cause interaction for this experiment was only marginally significant, $\underline{F}$ (2,117) = 2.47, $\underline{p} < .09$. Cell means for this interaction appear in Table 14. The Duncan test indicated that the significant main effect for valence was not

demonstrated in the inhibitory-cause condition. Thus, it appears that the facilitative-cause and control conditions led to a comparable decrease in attributions of responsibility in the good-outcome condition. Moreover, it should be noted that this pattern of results for the Valence X Cause interaction generally held for all grades (see Table 15). Contrary to predictions, there is no evidence that eighth-grade subjects would discount responsibility when consensus information is facilitative to either the good or bad outcome.

TABLE 14

Mean Attribution of Responsibility According to Valence of Outcome and External Cause (Consensus)

	External Cause		
Valence	Facilitatory	Inhibitory	Control
Bad	3.56_{a}	3.50_{ab}	3.62_{a}
Good	3.10_{c}	3.49_{ab}	3.21_{bc}

Note. Higher means indicate higher attributions of responsibility.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

There was also a modestly significant Grades X Cause X Severity interaction, F (4,117) = 2.61, $p < .05$. While this

TABLE 15

Mean Attributions of Responsibility According to Valence of Outcome, Grade, and External Cause (Consensus)

Grade	External Cause		
	Facilitative	Inhibitory	Control
	Bad Outcome		
2	3.46_{abcd}	3.46_{abcd}	3.61_{bcd}
5	3.96_{d}	3.68_{bcd}	3.79_{cd}
8	3.25_{abc}	3.36_{abcd}	3.46_{abcd}
	Good Outcome		
2	3.21_{abc}	3.54_{abcd}	3.14_{ab}
5	3.14_{ab}	3.79_{cd}	3.25_{abc}
8	2.93_{a}	3.14_{ab}	3.25_{abc}

Note. Higher means indicate higher attributions of responsibility.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

interaction does not lend itself easily to interpretations, it does indicate that the main effect for severity on attributions

of responsibility was not demonstrated consistently by all grades-external cause conditions (see Figure 4). For example, control-group eighth-grade children demonstrated a salient simple effect for severity which was much less evidenced by second- and fifth-grade controls. This particular finding was contrary to expectations.

The main effect for cause was not significant, nor were the remaining interactions.

Diagnosticity. Analysis of variance on "how often" scores (see Appendix 14) revealed a highly significant main effect for valence, F (1,117) = 66.94, $p < .01$, with more diagnosticity attributed to good acts (M = 2.82) than to bad acts (M = 2.45). However, there was also a significant Grades X Valence interaction, F (2,117) = 5.51, $p < .01$. Cell means for this interaction (see Table 16) indicated that the valence main effect was particularly salient with fifth-grade subjects, although simple effects for valence were significant for second- and eighth-grade subjects as well.

It had been predicted that the Grades X Cause interaction would indicate that eighth-grade subjects in the facilitative-cause condition would perceive less diagnosticity in actors' behavior than would their no-information counterparts. This interaction did not approach significance, but as the cell means indicate (see Figure 5), the difference between the respective means are in the predicted direction. A t-test on this difference provided marginal support for the hypothesis, t (117)

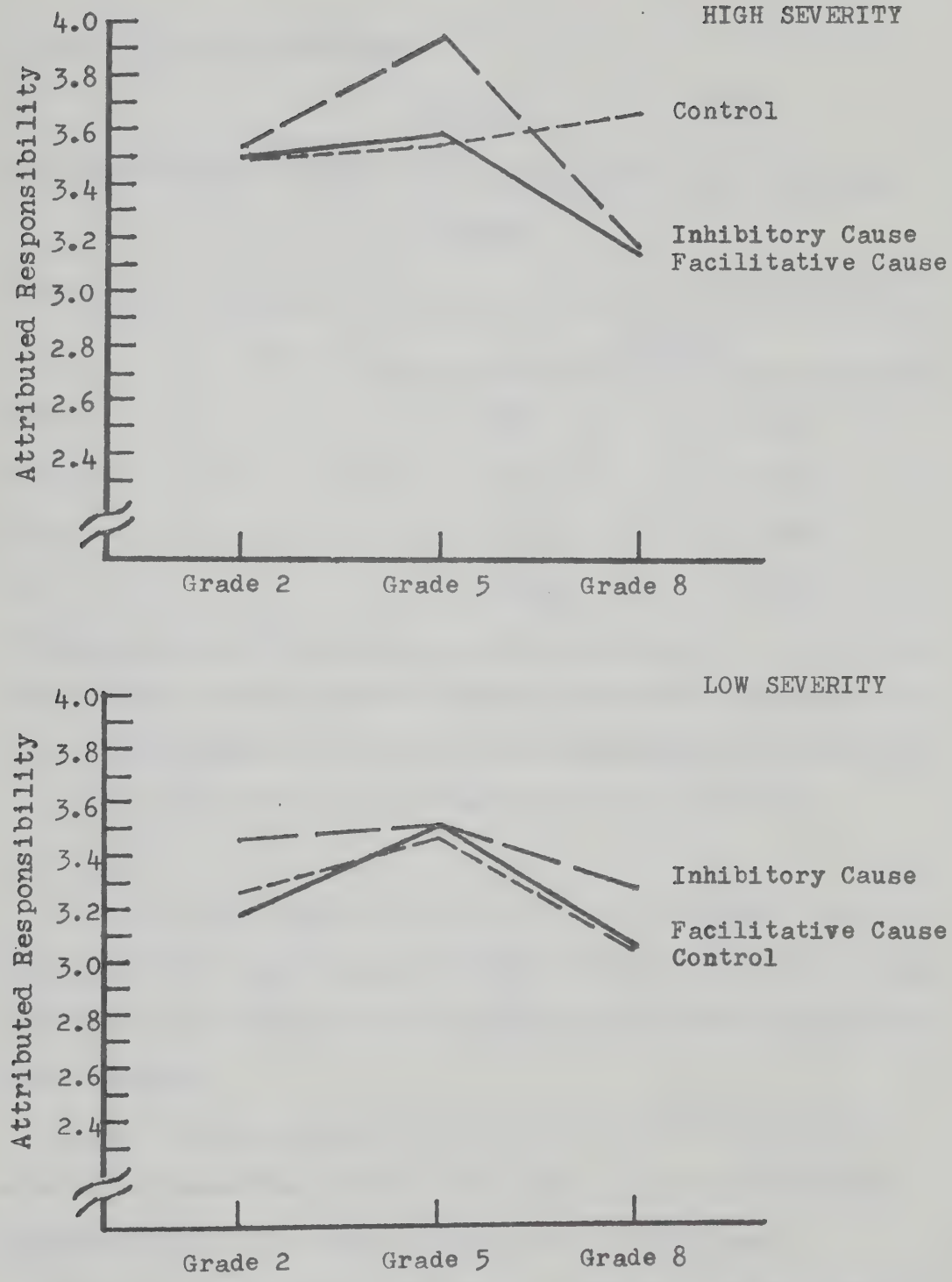

Figure 4. Mean Attributions of Responsibility According to Grades, Severity, and External Cause.

= 1.60, $p < .06$, one-tailed.

TABLE 16

Mean Diagnosticity Scores According to Grades and Valence of Outcome

	Valence	
Grade	Bad	Good
2	2.35_{a}	2.91_{c}
5	2.08_{b}	2.94_{c}
8	2.31_{ab}	2.60_{d}

Note. Higher means indicate higher attributed diagnosticity. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

Figure 5 also indicates that there is no support for the hypothesis that eighth-grade subjects in the inhibitory-cause condition would perceive more diagnosticity than would control eighth-graders.

There was also a significant Grades X Cause X Valence X Severity interaction, F (4,117) = 2.45, $p < .05$, which was uninterpretable. No other effects were significant.

Allocated Punishment. Analysis of variance on this measure (see Appendix 15) revealed three significant findings. First, as in Experiment 1, there was a highly significant main effect

for severity, $\underline{F}$ (1,117) = 246.9, $\underline{p} < .01$, with more punishment allocated to the actor involved in the high-severe act (stealing a bicycle; M = 3.71) than the low-severe act (stealing a pencil; M = 2.56). Again, there was also a main effect for grades, $\underline{F}$ (2,117) = 9.97, $\underline{p} < .01$. Eighth-grade subjects allocated least punishment overall (M = 2.83), while second- and fifth-grade subjects allocated comparable levels of punishment (Ms = 3.27 and 3.29, respectively). However, a significant Grades X Severity interaction, $\underline{F}$ (2,117) = 5.37, $\underline{p} < .01$, indicated that the grades main effect was particularly salient in the low-severe condition, and much less so in the high-severe condition (see Table 17).

TABLE 17

Mean Allocated Punishment According to Grades and Severity of Outcome

	Severity	
Grade	High	Low
2	3.81_{a}	2.74_{c}
5	3.74_{ab}	2.83_{c}
8	3.57_{b}	2.10_{d}

Note. Higher means indicate higher levels of allocated punishment.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

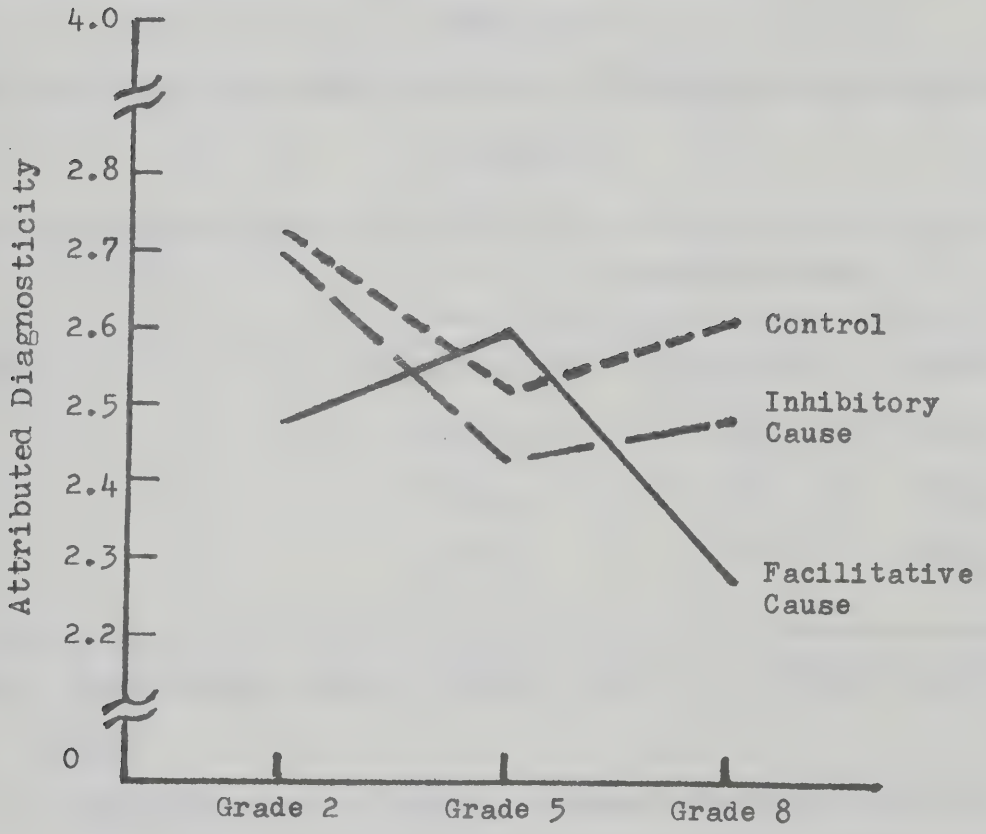

Figure 5. Mean Diagnosticity Scores According to Grades and External Cause (Consensus).

While the Grades X Cause interaction for this measure did not reach significance, the Duncan test on cell means (see Table 18) again provides support for the hypothesis that eighth-grade subjects would allocate less punishment in the facilitative-cause condition than would control subjects.

TABLE 18

Mean Allocated Punishment According to Grades and External Cause (Consensus)

	External Cause		
Grade	Facilitative	Inhibitory	Control
2	3.18_{abc}	3.21_{bc}	3.43_{c}
5	3.29_{bc}	3.36_{c}	3.21_{bc}
8	2.57_{d}	3.04_{ab}	2.89_{a}

Note. Higher means indicate higher levels of allocated punishment.

Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

There were no other significant effects in this analysis.

Allocated Reward. Analysis of variance on scores for allocated reward for good acts (see Appendix 16) indicated the same significant effects as found for allocated punishment. First, a significant main effect for grades, F (2,117) = 5.32, $p < .01$, revealed that eighth-grade subjects had allocated least reward

overall (M = 2.42), while second-graders had allocated the most (M = 3.05). Fifth-grade subjects scored intermediately (M = 2.82). Also, the main effect for severity, $\underline{F}$ (1,117) = 47.3, $\underline{p} < .01$, indicated that more reward was allocated for the high-severe act (M = 3.09) than for the low-severe act (M = 2.44). The significant Grades X Severity interaction, $\underline{F}$ (2,117) = 3.51, $\underline{p} < .05$, did not substantially alter the pattern of either main effect (see Table 19).

TABLE 19

Mean Allocated Reward According to Grades and Severity of Outcome

	Severity	
Grade	High	Low
2	3.50_{a}	2.60_{c}
5	2.98_{b}	2.67_{c}
8	2.79_{bc}	2.05_{d}

Note. Higher means indicate higher levels of allocated reward. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

The predicted Grades X Cause interaction did not approach significance, nor were specific predictions regarding this interaction supported by further statistical comparisons of cell means (see Table 20).

TABLE 20

Mean Allocated Reward According to Grades and External Cause (Consensus)

	External Cause		
Grade	Facilitative	Inhibitory	Control
2	3.04_{ab}	3.07_{ab}	3.04_{ab}
5	2.82_{abc}	3.14_{a}	2.50_{c}
8	2.32_{c}	2.36_{c}	2.57_{bc}

Note. Higher means indicate higher levels of allocated reward. Means which do not share a common subscript differ at the .05 level by the Duncan multiple range test.

Discussion

Again, the valence-of-outcome effect was obtained in Experiment 2, with more responsibility attributed to actors involved in bad acts than in good acts. However, unlike Experiment 1, there was no indication that older children would demonstrate this effect any less than younger children. In fact, the valence main effect in Experiment 2 was salient in all external-cause conditions and for all grades, and reached a high level of statistical significance.

Also replicated in this study was the reversal of the valence main effect on diagnosticity scores. Actors involved in good acts were perceived as more likely to repeat similar

behaviors than were bad-outcome actors, despite the fact that the former were held less responsible for the good outcomes.

Generally, it can be seen that the external cause manipulation in this experiment, i.e., consensus information, did not have a strong effect upon any of the dependent variables. However, there are some exceptions to this generalization. First, the Valence X Cause interaction on attributions of responsibility, while itself only marginally significant, provided some support for the hypothesis that, for good outcomes, facilitative-cause information and no information would be treated comparably, with relatively more responsibility attributed by subjects exposed to consensus information which is inhibitory to the good outcome. On the other hand, contrary to predictions, this apparent "augmenting" of responsibility by subjects in the latter condition was not more salient with eighth-grade children. In fact, further examination revealed that these predicted differences were obtained solely by second- and fifth-grade subjects, particularly the latter. The possibility that younger children were in fact augmenting responsibility in the inhibitory-cause/good-outcome condition is further supported by data on allocated reward, at least in the case of fifth-graders, who perceived more reward-deservingness in the inhibitory-cause condition than in the control condition. This unexpected pattern of results would seem to suggest that it was fifth-grade children and not eighth-graders to whom inhibitory-consensus information served as a salient external cause.

However, there is some indication that eighth-grade children did behave in the predicted fashion when the consensus information was facilitative. Marginally less diagnosticity was attributed by eighth-grade children in the facilitative-cause condition than in the control condition, regardless of valence of outcome. A parallel difference in the predicted direction on attributions of responsibility did not approach significance. Moreover, an interesting pattern of results occurred on scores for allocated punishment, which was consistent with predictions. Although eighth-grade children were significantly less punitive overall than were younger subjects, within-grades comparisons revealed that eighth-graders, and only eighth-graders, were yet more disinclined to allocate punishment in the facilitative-cause condition than were control-group eighth-graders.

Nevertheless, this significant finding on allocated punishment, in the absence of support from parallel comparisons on attributions of responsibility, and only marginal support on the diagnosticity measure, suggests that some alternative explanation to decreased attributed responsibility in the facilitative-cause condition may be more viable. Indeed, it appears that eighth-grade subjects may have perceived the observation that "everybody does it" as a relatively weak excuse for performing bad acts, but nevertheless a good reason for being less punitive toward such actors. In this light, it is unfortunate that children's moral judgments were not collected in this study. For example,it would have been interesting to discover whether

or not eighth-grade children in the facilitative-cause/bad-outcome condition were also judging actors to be less "bad". The hypothesis that a moral judgment mediates attributions of responsibility and subsequent sanctioning, which has been suggested in this thesis, would predict such a decrease in judgments of bad. On the other hand, the apparent "amnesty" effect obtained in the data does not necessitate a mitigated moral judgment. Future research on the role of consensus information on children's moral judgments could shed light on this question.

Finally, predictions regarding the severity of outcomes were not well supported. While a highly significant main effect for severity on attributions of responsibility did confirm the hypothesis that more responsibility would be attributed for high- than for low-severe acts, the subsidiary hypothesis that this effect would be more salient in younger children was not supported. Also unsupported was the expectation that high-severe acts would lead to higher diagnosticity. This lack of a significant main effect for severity on diagnosticity scores is especially interesting, considering the fact that actors were held more responsible for high-severe acts than for low-severe acts. It is also noteworthy that there was almost a total absence of significant interactions between severity and external cause on all of the dependent variables, which is in sharp contrast to the results of Experiment 1.

General Discussion

A central purpose of this thesis was to investigate Kelley's (1971b) provocative attempt to subsume moral judgments under the propositions of attribution theory, particularly those propositions pertaining to attributions of responsibility. "The moral goodness of a person," wrote Kelley, "is a matter of person attribution and is judged according to the same criteria as are used in other judgments of personal causation or responsibility" (p. 294).

However, the results of both experiments reported here, together with a growing literature on extra-theoretical "biases" in attributions of responsibility (cf. Shaver, 1975, chp. 7), caution against an uncritical application of a theory as information-processing-oriented as Kelley's, to realms of behavior as affect-laden as moral judgments and attributions of responsibility. To be fair, Kelley himself admitted that attribution theory may provide only a "partial base" for moral judgments. It would seem that future empirical work on these issues would best address itself to specifying what this "base" does and does not include, and moreover how it interacts with affective-motivational systems in the production of moral judgments.

The problem is further complicated when we consider the development of responsibility attributions, as in this thesis. One conclusion that can be derived fairly readily from these results is that moral judgments and attributions of responsibility become more interdependent with development, such that

judgments of good and bad coincide remarkably well with judgments of responsibility in eighth-grade children, and hardly at all in second-grade children. For example, only the latter subjects allocated severe punishments to actors for bad outcomes for which they were not held responsible.

Nevertheless, the propositions of attribution theory (and derivations thereof) were upheld in a number of significant findings. Perhaps the best-supported hypothesis in this regard was the predicted main effect for valence on attributions of responsibility, a finding which has occurred in previous research but has received little attention. Both experiments, particularly Experiment 1, provided evidence that the valence-of-outcome effect is a function of different assumptions made by perceivers when judging good versus bad acts. That is, causes external to the actor are assumed to be facilitative to the performance of good acts, and inhibitory to the performance of bad acts. These different assumptions therefore lead to increased internal responsibility attributed to actors involved in the latter.

The results of Experiment 1, compared to those of Experiment 2, attest to the relatively serious consideration that children give to perceived choice when making causal attributions. In the bad-outcome condition, even second-grade subjects seemed to discount the responsibility of story characters who performed such acts after being ordered to do so. It is interesting to

compare these results to those of Costanzo, Grumet and Brehem (1974) and Smith (1975), two studies which have indicated that children may not begin discounting the intrinsic motivation of story-characters who perform relatively neutral acts in the presance of adult constraint, until approximately the fourth grade. Thus, it is possible that the discounting principle may operate sooner in the child's life when the attribution involves a moral evaluation or a responsibility attribution than when the attribution is morally neutral in the child's eye. It should be noted that the second-grade children in the Shaw and Sulzer (1964, Experiment 1) study also attributed less responsibility to story characters involved in bad outcomes facilitated by an adult (Justified Commission Level). Unfortunately, Shaw and Sulzer did not report a statistical test of this difference.

However, the results of Experiment 1 provide further insight into the developmental pattern of the attribution process, beyond the finding that even second-graders will discount internal responsibility in the presence of an external facilitative cause. When the parental dictate was <u>inhibitory</u> to the performance of good acts, fifth- and eighth-graders--but not second-graders--attributed <u>increased</u> internal responsibility. These findings are notably consistent with those reported by Shultz and Mendelson (1975) on children's perceptions of physical causality. In that study, even children in the youngest age group (3 - 4 years) identified facilitative causes significantly more often than inhibitory causes. This was also the case with the

next older age group (6 - 7 years). Only children in the oldest age group (9 - 11 years) identified both facilitative and inhibitory causes with comparable success. Thus, it is apparent that in the areas of both social and physical causality, children learn to identify facilitative causes before inhibitory causes.

Also relevant to development trends are the findings for diagnosticity. The results of both experiments indicated that eighth-grade children--and only eighth-graders--discounted the diagnosticity of actions performed in the presence of an external facilitative cause. However, not even eighth-grade children perceived greater diagnosticity for acts performed in the presence of an inhibitory cause. It may be that attribution theorists who have dealt specifically with diagnosticity-like phenomena (e.g., Brickman, Ryan & Wortman, 1975; Jones & Davis, 1965; Trope & Burnstein, 1975) have been discussing something that emerges in the formal operations period of Piaget's cognitive developmental hierarchy. That is, diagnosticity is in effect an inference based on a prior inference (i.e., responsibility), and therefore may require the relatively abstract cognitive development acquired in early adolescence. A comprehensive theory of the development of attribution processes would have to take into account the sheer cognitive difficulty involved in making diagnosticity judgments, as well as experiential factors which facilitate such interpersonal perceptions.

An expectation that was not well supported by the experimental

results was the hypothesis that consensus serves as a potent external cause in attributions of responsibility (cf. Kelley, 1971b; Ross & DiTecco, 1975). Of course, it is possible that the children involved in these experiments were too young to regard consensus information as having any significance for social behavior. However, all relevant research militates against this possibility. For example, conformity studies with children have shown that conformity to the perceptual judgments of peers is highest at approximately 12 years of age (Costanzo & Shaw, 1966), or possibly even younger still (cf. Berenda, 1950). Clearly, the influential effect of peer consensus is not strictly an adult phenomenon.

What is likely, however, is that consensus information works more subtly than has been supposed by researchers on attribution theory. First, it may be extremely difficult, if not impossible, to manipulate consensus effectively when subjects are asked to make a moral evaluation, or any judgment which may elicit subjects' existing beliefs about "what most people do." Thus, telling children that, e.g., stealing bicycles is a common occurrence in a story character's school may run counter to beliefs regarding consensus which the children bring to the experimental situation. This possibility is supported by the manipulation check data of Experiment 2, which indicated that all children did in fact maintain beliefs about the good and bad behavior of "other kids", over and above what they were told by the experimenter.

While there can be little doubt that such factors as the perceptions and opinions of the majority can have powerful effects on the conformity of individuals, it nevertheless remains possible that individuals do not spontaneously regard consensus as an important variable when making a causal attribution. A finding reported rather incidentially by Milgram (1965) is instructive in this regard. Milgram compared the shock-emitting behaviors of individuals participating alone or with two experimental confederates who refused to shock the "learner" in the ubiquitous Buss-Milgram paradigm. As usual, subjects participating alone very frequently yielded to the experimenter's demand to deliver high levels of shock to the learner. However, subjects who witnessed the defiance of two fellow subjects seldom delivered more than a trivial amount of shock before similarly refusing to obey the experimenter. What is interesting about this study are the subsequent verbalizations given by the latter regarding the reasons for their own defiance. Seventy-five per cent of the defiant subjects clearly insisted that they would have refused the experimenter's demand for higher shock without the confederates' example. Moreover, only four of 36 defiant subjects admitted that the confederates' rebellion had any causal role in their own defiant behavior.

The point to be made from this example is that, while consensus may cause changes in individual behavior, consensus information may have little effect on the <u>perceptions</u> of causality of individual behavior. This distinction may be particularly true

in experiments utilizing hypothetical situations (e.g., stories) and paper-and-pencil measures. On the other hand, it may be that consensus information could play a role in attributions when individuals are tempted to participate in deviant behavior themselves, as was suggested by Ross and DiTecco (1975), or when they need to rationalize such behavior which they have already committed. These are interesting possibilities and deserving of future research consideration.

A discussion of the development of causal attributions would not be complete without some consideration of the literature on locus of control, a personality dimension which has many conceptual similarities with attribution theory (cf. Joe, 1974). Briefly, locus of control refers to generalized expectancies that one's reinforcements are controlled by the self (internal) or are controlled by factors external to the self (Lefcourt, 1966; Rotter, 1966). Such expectancies are presumed to be a function of one's learning history, such that a person who has had considerable experience with reinforcement being contingent on his own behavior would develop an expectancy that future reinforcements would similarly be internally controlled; whereas a person whose reinforcements have not usually been contingent on his behavior develops an expectancy that future reinforcements will normally be the result of powerful others, luck and fate.

Research on the development of internal vs. external control has been generally consistent in suggesting that the parents of

"internals" tend to be more permissive, less rejecting, more consistent in discipline, and more encouraging of independence and achievement, than the parents of "externals" (e.g., Davis & Phares, 1969; Katkovsky, Crandall & Good, 1967). Also important is the finding that the internal orientation tends to increase during middle childhood, possible due to the child's increasing competence in controlling his own outcome (Bialer, 1961).

The link between locus of control and attributions of responsibility is provided by the assumption that people project their control orientations when making causal attributions regarding the behavior of others (Sosis, 1974). This reasoning is similar to Heider's (1958) proposal that people assume that conditions operating upon them must also operate upon others. Thus, Sosis (1974) found support for her hypothesis that internal-oriented adults would attribute more personal responsibility to others involved in accidents and would judge them more harshly than would externals, while the latter were more likely to attribute the accidents to bad luck.

No study has directly examined the locus-of-control/attribution-of-responsibility relationship in a developmental context. However, a very recent study by Karlovac, Feldman, Higgins and Ruble (Note 2) claimed to provide indirect support for a developing positive relationship between internal locus-of-control and tendencies to attribute internal causality to others. Subjects

in the Karlovac et al. study were pre-schoolers, third and fourth-graders, and adults who made person (internal) vs. entity (external) attributions of causality for the stimulus preferences of a target person. The results showed a main effect for age, with the adults making significantly more person attributions in all experimental conditions than either of the two groups of children. Differences between the latter groups were specific to some experimental conditions.

However, the present research fails to provide support for the hypothesis that persons attribute more personal responsibility with increasing age. In fact, both experiments yielded a curvilinear relationship between age and attributed responsibility, with fifth-grade subjects attributing most internal responsibility overall compared to second- and eighth-graders. The reason for this distinct pattern of results is unknown. It is difficult to compare these data to the Karlovac et al. results for two reasons: the ages of subjects do not coincide, and the types of attributions required of subjects differ. Future research on these issues would do well to compare developmental trends in attributions of moral responsibility with trends in attributions of internal versus external causality.

To conclude this discussion, it is appropriate to reiterate that a very few number of situations were presented to the children, and generalizations suggested from the experiments are severely restricted until this research cam be replicated with new situations assumed to operationalize the same indepen-

dent variables (valence, severity, cause). Moreover, this precaution is underscored by the probability, alluded to earlier, that seemingly irrelevant details of situations will sometimes have specific and extra-theoretical effects on attributions of responsibility and moral evaluations. For example, one fifth-grade boy, before he could decide on the amount of responsibility to be attributed to the bicycle thief, wanted to know if the bicycle was a new one or an old one.

Footnotes

[1]The use of "consensus" in this dissertation refers to consensus as a basis for normative behavior, and not as an information-processing variable used for validating causal attributions. Kelley has used consensus in both contexts. The latter use of consensus was introduced by Kelley in his ANOVA model of attribution theory as one of three criteria for judging the validity of one's attributions (the other two criteria being the distinctiveness of an event and its consistency over time; cf. Kelley, 1973). In this context, high consensus means that there is similarity in response to a stimulus across persons, and therefore there must be something about the stimulus that causes the event. However, in his attempt to explain moral attributions, Kelley (1971b) linked consensus to Heider's (1958) concept of "ought", which, although being an abstract standard, was presumed by Heider to be experienced by persons as having an objective reality which determines individual behavior. Ross and DiTecco (1975) used consensus in a similar fashion in their discussion of "oughts as impersonal, consensually-defined standards." In this context, consensus acquires a directly causal role, since it determines the perceived standard for appropriate behavior.

The differing roles of consensus in moral vs. non-moral attributions can be illustrated by considering Kelley's (1971b) description of the dilemma involved in attributing responsibility to individuals for good acts. The "normative" consensus

(as it is used by Kelley, 1971b) makes for difficulty in assigning praiseworthiness to persons who perform good acts that conform to a norm: the norm itself is usually seen as sufficient cause for the performance of the act. However, the high consensus of the ANOVA model would suggest that people would probably infer that there was something about the object of the good act that causes people to perform the act, rather than consensus per se. Consensus in this latter context does not have causal value.

The issue seems to be that when there is an evaluative dimension to the attribution--i.e., when the goodness-badness dimension is salient--consensus information is transformed from a method of validating causality to the very basis by which the behavior is to be judged.

[2]This rather extended definition of responsibility was an attempt to improve on that used by Shaw and Sulzer (1964), which merely equated responsibility for bad acts with blame, and responsibility for good acts with gratitude. Besides seeming too cursory, this latter definition appears to confound attributed responsibility with sanctioning.

[3]Relevant studies which explicitly stated an absence of sex differences in children's moral judgments include: Chandler, Greenspan and Barenboim, 1973; Gutkin, 1972; Hebble, 1971; Hill and Enzle, Note 1; Jenson and Hughston, 1971. Similar examples from the child-attribution literature include Baldwin and

Baldwin, 1970, and Shaw and Sulzer, 1964. Many studies in both areas included sex as a factor in the initial design, but did not mention it again while reporting their results.

Reference Notes

1. Hill, K. A., & Enzle, M. E. Interactive effects of training modality and age on children's moral judgments. Unpublished manuscript, University of Alberta, 1976.

2. Karlovac, M., Feldman, N. S., Higgins, E. T., & Ruble, D. N. The effect of consensus information on causal attributions. Paper presented at the annual convention of the American Psychological Association, Washington, D. C., September, 1976.

References

Baldwin, A. L., Baldwin, C. P., Castillo-Vales, V., & Seegmiller, B. Cross-cultural similarities in the development of the concept of kindness. In W. W.Lambert & R. Weisbrod (Eds.), Comparative perspectives on social psychology. Boston: Little, Brown, 1971.

Baldwin, C. P., & Baldwin, A. L. Children's judgments of kindness. Child Development, 1970, 41, 29-47.

Berenda, R. W. The influence of the group on the judgments of children. New York: King's Crown Press, 1950.

Berg-Cross, L. G. Intentionality, degree of damage, and moral judgments. Child Development, 1975, 46, 970-974.

Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. Journal of Personality, 1961, 29, 303-320.

Brickman, P., Ryan, K., & Wortman, C. B. Causal chains: Attributions of responsibility as a function of immediate and prior causes. Journal of Personality and Social Psychology, 1975, 32, 1060-1067.

Chandler, M. J., Greenspan, S., & Barenboim, C. Judgments of intentionality in response to videotaped and verbally presented moral dilemmas: The medium is the message. Child Development, 1973, 44, 315-320.

Costanzo, P. R., Grumet, J. F., & Brehem, S. S. The effects of choice and source of constraint on children's attributions

of preference. Journal of Experimental Social Psychology, 1974, 10, 352-364.

Costanzo, P. R., & Shaw, M. E. Conformity as a function of age level. Child Development, 1966, 37, 967-975.

Cunningham, J. D., & Kelley, H. H. Causal attributions for interpersonal events of varying magnitude. Journal of Personality, 1975, 43, 74-93.

Davis, L., & Phares, E. Parental antecedents of internal-external control of reinforcement. Psychological Reports, 1969, 24, 427-436.

Joe, V. C. Perceived personal control and attributions of causality. Perceptual and Motor Skills, 1974, 38, 323-329.

Hartup, W. W. Peer interaction and social organization. In P. Mussen (Ed.), Carmichael's manuel of child psychology, (Vol. 2). New York: Wiley, 1970.

Hebble, P. W. The development of elementary school children's judgment of intent. Child Development, 1971, 42, 1203-1215.

Heider, F. The psychology of interpersonal relations. New York: Wiley, 1958.

Jensen, L. C., & Hughston, K. The relationship between type of sanction, story content, and children's judgments which are independent of sanction. Journal of Genetic Psychology, 1973, 122, 49-54.

Jones, E. E., & Davis, K. E. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), Advances in experimental social psychology, (Vol. 2).

New York: Academic Press, 1965.

Katkovsky, W., Crandall, V. C., & Good, S. Parental antecedents of children's beliefs in internal-external control of reinforcements in intellectual achievement situations. Child Development, 1967, 38, 765-776.

Kelley, H. H. Attribution in social interaction. Morristown, N. J.: General Learning Press, 1971. (a)

Kelley, H. H. Moral evaluation. American Psychologist, 1971, 26, 293-300. (b)

Kelley, H. H. The process of causal attribution. American Psychologist, 1973, 28, 107-128.

Lefcourt, H. M. Internal versus external control of reinforcement: A review. Psychological Bulletin, 1966, 65, 206-220.

McKechnie, R. J. Between Piaget's stages: A study in moral development. British Journal of Educational Psychology, 1971, 41, 213-217.

Milgram, S. Liberating effects of group pressure. Journal of Personality and Social Psychology, 1965, 1, 127-134.

Piaget, J. The moral judgment of the child. New York: Harcourt, Brace, 1932.

Ross, M. & DiTecco, D. An attributional analysis of moral judgment. Journal of Social Issues, 1975.

Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80, No. 1 (Whole No. 609).

Schneider, F. W., & Shaw, M. E. Sanctioning behavior in Negro

and in white populations. Journal of Social Psychology, 1970, 81, 63-72.

Shaver, K. G. An introduction to attribution processes. Cambridge, Mass.: Winthrop, 1975.

Shaw, J. I., & Skolnick, P. Attribution of responsibility for a happy accident. Journal of Personality and Social Psychology, 1971, 18, 380-383.

Shaw, M. E., & Reitan, H. T. Attribution of responsibility as a basis for sanctioning behavior. British Journal of Social and Clinical Psychology, 1969, 8, 217-226.

Shaw, M. E., & Schneider, F. W. Intellectual competence as a variable in attributions of responsibility and assignment of sanctions. Journal of Social Psychology, 1968, 78, 31-39.

Shaw, M. E., & Sulzer, J. L. An empirical test of Heider's levels in attributions of responsibility. Journal of Abnormal and Social Psychology, 1964, 69, 39-46.

Shultz, T., Butkowsky, I., Pearce, J., & Shanfield, H. Development of schemes for the attribution of multiple psychological causes. Developmental Psychology, 1975, 11, 502-510.

Shultz, T., & Mendelson, R. The use of covariation as a principle of causal analysis. Child Development, 1975, 46, 394-399.

Smith, M. C. Children's use of the multiple sufficient cause schema in social perception. Journal of Personality and Social Psychology, 1975, 32, 737-747.

Sosis, R. H. Internal-external control and the perception of

responsibility of another for an accident. Journal of Personality and Social Psychology, 1974, 30, 393-399.

Strickland, L. H. Surveillance and trust. Journal of Personality, 1958, 26, 200-215.

Trope, Y., & Burnstein, E. Processing the information contained in another's behavior. Journal of Experimental Social Psychology, 1975, 11, 439-458.

Weiner, B., & Peter, N. A cognitive-developmental analysis of achievement and moral judgments. Developmental Psychology, 1973, 9, 290-309.

Wright, D. The psychology of moral behaviour. Baltimore: Penguin, 1971.

Appendix 1

Pilot Study

Method

Subjects. Thirty children (21 boys and 9 girls) were recruited from an Edmonton public school. Ages of children ranged from 6 years, 9 months to 12 years, 2 months. Mean age overall was 10 years. Children were divided evenly into two groups according to age. Mean age of children in the younger group was 8 years, 3 months. Mean age of children in the older group was 11 years, 8 months. All children were white and middle-class.

Materials. A total of 12 "consequences" were constructed for presentation to subjects. Six of the acts involved bad (antisocial) behavior and six involved good (prosocial) behavior. The only other criterion for selection was that acts vary in severity.

The bad-consequence acts were:

1. Someone taking a pencil that doesn't belong to them.
2. Someone hitting someone else hard on the nose and making it bleed, on purpose.
3. Someone telling a fib to a friend so they won't have to play with them.
4. Someone taking a bicycle that doesn't belong to them.
5. Someone knocking someone else's book out of their hand, on purpose.
6. Someone telling a fib to the teacher to stay home from school.

Appendix 1 (continued)

The good-consequence acts were:

1. Someone giving someone else directions on how to get to the grocery store.

2. Someone climbing a tree to help a child get back down.

3. Someone sharing their lunch with a friend.

4. Someone helping a child on the playground tie their shoes.

5. Someone giving all their savings, which is a lot of money, to poor children.

6. Someone helping a child in the park find their mittens.

Scales. For purposes of rating the "badness" of bad acts and the "goodness" of good acts, two 4-point scales were devised. For judgments of badness, the scale ranged from (1) "not bad or good", to (2) "a little bit bad", to (3) "pretty bad", to (4) "very bad". Similarly, the scale for judgments of goodness ranged from (1) "not good or bad", to (2) "a little bit good", to (3) "pretty good", to (4) "very good." Accompanying each scale value was a simple illustration of a face of a devil or an angel, according to the valence of the acts being judged. For badness judgments, values of 2, 3 and 4 were positively correlated with the size of the devil's horns. These values for the goodness scale were positively correlated with the size of the angel's halo. The face accompanying value 1 had neither halo nor horns and was identical for both scales. These illustrations were included in order to facilitate comprehension of

Appendix 1 (continued)

the scales by younger children.

Procedure. Each subject was tested individually by a female experimenter. Instructions to the child were as follows:

> Sometimes people do bad (good) things, but some bad (good) things are worse (better) than others, and other things people do are hardly bad (good) at all. I am going to tell you about some things people sometimes do, and I would like you to tell me how bad (good) you think it is to do these things.

The experimenter explained the use of the rating scales to the child, then proceeded to read the items. After each item was read aloud, the child indicated his response by pointing to the appropriate scale value. Presentation of bad and good outcomes were counterbalanced.

Results

The mean and standard deviation for each item was computed separately for younger and older children. These results appear in Table 21. Numeration of bad and good acts refers to the numeration presented in the Method section.

On the basis of these data, items one and four from the list of bad acts were selected as the mild and severe outcomes, respectively. Two-tailed t-tests indicated that both younger and older children judged item four to be significantly more

Appendix 1 (continued)

severe than item one, ts (13) $>$ 4.5, ps $<$ .01. The differences between age groups for both the severe and mild items did not reach statistical significance (ps $>$ .05).

For the good-act items, item four was selected as the mild outcome and item five as the severe. Two-tailed t-tests indicated that both younger and older children judged item five to be significantly more "good" than item four, ts (13) $>$ 3.75, ps $<$ .01. Also, the differences between age groups on both the mild and severe "good" items did not reach statistical significance (ps $>$ .10).

There was a tendency for children of both age groups, particularly the younger children, to rate the mild good act as slightly more severe than the mild bad act, thereby suggesting a possible positive bias on the mild-outcome items. However, these differences approached only marginal significance, even with the younger children, t (13) = 2.07, $p > .05$. Therefore, these items were retained.

Appendix 1 (continued)

TABLE 21

Means and Standard Deviations for Judgments of Bad and Good Acts by Younger and Older Children

		Younger Children		Older Children	
		M	sd	M	sd
Bad Acts	1	2.07	.917	2.07	.475
	2	3.36	.497	3.50	.519
	3	2.29	.994	2.71	1.204
	4	3.64	.497	4.00	.000
	5	2.63	.745	2.64	.633
	6	3.07	1.141	3.43	.646
Good Acts	1	3.00	.784	2.79	.579
	2	3.50	.855	3.71	.469
	3	3.14	.864	2.93	.730
	4.	2.93	.730	2.43	.852
	5	3.93	.267	4.00	.000
	6	3.07	.616	2.86	.363

Appendix 2

Description of Stimulus Materials for Experiment 1

Bad-Outcome/High-Severe Story

Facilitative-cause condition. "Bobby's mother told Bobby to take someone else's bicycle and keep it for himself. So Bobby took another boy's bicycle and kept it."

Inhibitory-cause condition. "Bobby's mother told him to never take bicycles that didn't belong to him. But Bobby took another boy's bicycle and kept it for himself."

Control Condition. "Bobby took another boy's bicycle and kept it for himself."

Dependent measures. A. "Afterwards, how responsible was Bobby for the other boy not having the bicycle anymore?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How bad was Bobby for taking the other boy's bicycle?" (1) not bad at all, (2) a little bit bad, (3) pretty bad, (4) very bad. C. "How much should Bobby be punished for taking the other boy's bicycle?" (1) not punished at all, (2) punished a little bit, (3) punished pretty much, (4) punished a lot. D. "How often does Bobby take other things besides bicycles that don't belong to him?" (1) not often at all, (2) once in awhile, (3) pretty often, (4) almost all of the time. E. "How mad was Bobby's mother when she found out that Bobby had taken the

Appendix 2 (continued)

other boy's bicycle?" (1) not mad at all, (2) a little bit mad, (3) pretty mad, (4) very mad.

Bad-Outcome/Low-Severe Story

Facilitative-cause condition. "John's mother told John to take someone else's pencil and keep it for himself. So John took another boy's pencil and kept it."

Inhibitory-cause condition. "John's mother told John to never take pencils that didn't belong to him. But John took another boy's pencil and kept it for himself."

Control condition. "John took another boy's pencil and kept it for himself."

Dependent measures. A. "Afterwards, how responsible was John for the other boy not having the pencil anymore?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How bad was John for taking the other boy's pencil?" (1) not bad at all, (2) a little bit bad, (3) pretty bad, (4) very bad. C. "How much should John be punished for taking the other boy's pencil?" (1) not punished at all, (2) punished a little bit, (3) punished pretty much, (4) punished a lot. D. "How often does John take other things besides pencils that don't belong to him?" (1) not often at all, (2) once in awhile, (3) pretty often, (4) almost all of the time. E. "How mad was John's mother when she found out that John had taken another boy's

Appendix 2 (continued)

pencil?" (1) not mad at all, (2) a little bit mad, (3) pretty mad, (4) very mad.

Good-Outcome/High-Severe Story

Facilitative-cause condition. "Steve's mother told Steve to give all of his savings to poor children. So Steve gave all of his savings, which was a lot of money, to poor children."

Inhibitory-cause condition. "Steve's mother told Steve not to give any of his savings to poor children. But Steve gave all of his savings, which was a lot of money, to poor children."

Control condition. "Steve gave all of his savings, which was a lot of money, to poor children."

Dependent measures. A. "Afterwards, how responsible was Steve for the poor children having all of his savings?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How good was Steve for giving his savings to poor children?" (1) not good at all, (2) a little bit good, (3) pretty good, (4) very good. C. "How much should Steve be rewarded for giving his savings to poor children?" (1) not rewarded at all, (2) rewarded a little bit, (3) rewarded pretty much, (4) rewarded a lot. D. "How often does Steve help other people in other ways besides giving them money?" (1) not often at all, (2) once in awhile, (3) pretty often, (4) almost all of the time.

Appendix 2 (continued)

E. "How happy was Steve's mother when she found out that Steve had given all his savings to poor children?" (1) not happy at all, (2) a little bit happy, (3) pretty happy, (4) very happy.

Good-outcome/Low-Severe Story

Facilitative-cause condition. "Andy's mother told Andy to help younger children tie their shoelaces. So Andy helped a little boy on the playground tie his shoelaces."

Inhibitory-cause condition. "Andy's mother told Andy not to help younger children tie their shoelaces. But Andy helped a little boy on the playground tie his shoelaces."

Control condition. "Andy helped a little boy on the playground tie his shoelaces."

Dependent measures. A. "Afterwards, how responsible was Andy for the little boy having his shoelaces tied?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How good was Andy for tying the little boy's shoelaces?" (1) not good at all, (2) a little bit good, (3) pretty good, (4) very good. C. "How much should Andy be rewarded for tying the little boy's shoelaces?" (1) not rewarded at all, (2) rewarded a little bit, (3) rewarded pretty much, (4) rewarded a lot. D. "How often does Andy help other people in other ways besides tying shoelaces?" (1) not often at all, (2) once in awhile, (3) pretty

Appendix 2 (continued)

often, (4) almost all of the time. E. "How happy was Andy's mother when she found out that Andy had helped a little boy on the playground tie his shoelaces?" (1) not happy at all, (2) a little bit happy, (3) pretty happy, (4) very happy.

Appendix 3

Summary of Analysis of Variance of Scores on "How Mad" Judgments (Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	14.91	7.45	6.80	<.002
B: Cause	2	87.11	43.56	39.76	<.001
A x B	4	2.74	.68	.62	
Error	108	118.31	1.10		
C: Severity	1	39.39	39.39	120.70	<.001
A x C	2	1.95	.97	2.99	<.06
B x C	2	11.39	5.69	17.45	<.001
A x B x C	4	2.05	.51	1.57	
Error	108	35.23	.33		

Appendix 4

Summary of Analysis of Variance of Scores

on "How Happy" Judgments

(Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	3.26	1.63	2.50	<.09
B: Cause	2	226.33	113.17	173.60	<.001
A x B	4	3.56	.89	1.37	
Error	108	70.39	.65		
C: Severity	1	2.26	2.26	6.31	<.02
A x C	2	.52	.26	.73	
B x C	2	6.01	3.00	8.39	<.001
A x B x C	4	1.02	.25	.71	
Error	108	38.69	.36		

Appendix 5

Summary of Analysis of Variance of Scores

on Judgments of Bad

(Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	2.03	1.01	1.74	
B: Cause	2	5.33	2.67	4.58	<.02
A x B	4	6.26	1.56	2.69	<.04
Error	108	62.92	.58		
C: Severity	1	15.39	15.39	42.02	<.001
A x C	2	.08	.04	.11	
B x C	2	2.87	1.44	3.92	<.03
A x B x C	4	1.13	.28	.77	
Error	108	39.54	.37		

Appendix 6

Summary of Analysis of Variance of Scores on Judgments of Good (Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	5.85	2.92	4.63	<.02
B: Cause	2	6.33	3.17	5.01	<.01
A x B	4	19.05	4.76	7.54	<.001
Error	108	68.23	.63		
C: Severity	1	3.35	3.35	10.97	<.002
A x C	2	.96	.48	1.57	
B x C	2	2.68	1.34	4.38	<.02
A x B x C	4	3.02	.75	2.47	<.05
Error	108	33.00	.31		

Appendix 7

Summary of Analysis of Variance of Scores on Attributions of Responsibility (Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	18.90	9.45	8.09	<.001
B: Cause	2	43.15	21.57	18.46	<.001
A x B	4	4.25	1.06	.91	
Error	108	126.19	1.17		
C: Valence	1	2.33	2.33	1.80	
A x C	2	.50	.25	.19	
B x C	2	13.28	6.64	5.14	<.01
A x B x C	4	7.06	1.77	1.37	
Error	108	139.58	1.29		
D: Severity	1	.62	.62	1.19	
A x D	2	1.85	.93	1.78	
B x D	2	7.91	3.96	7.63	<.001
A x B x D	4	5.33	1.33	2.57	<.05
Error	108	56.04	.52		
C x D	1	.10	.10	.23	
A x C x D	2	1.06	.53	1.16	
B x C x D	2	.48	.24	.53	

Source	df	SS	MS	F	p
A x B x C x D	4	1.00	.25	.55	
Error	108	49.12	.45		

Appendix 8

Summary of Analysis of Variance of Scores on Attribution of Diagnosticity (Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	.81	.40	.37	
B: Cause	2	5.40	2.70	2.46	<.10
A x B	4	7.29	1.82	1.66	
Error	108	118.42	1.10		
C: Valence	1	5.78	5.78	7.59	<.01
A x C	2	.004	.002	.003	
B x C	2	.90	.45	.59	
A x B x C	4	3.12	.78	1.03	
Error	108	82.19	.76		
D: Severity	1	.01	.01	.03	
A x D	2	1.82	.91	2.70	<.08
B x D	2	.93	.46	1.37	
A x B x D	4	2.74	.68	2.03	<.10
Error	108	36.50	.34		
C x D	1	1.44	1.44	5.03	<.03
A x C x D	2	.85	.43	1.48	
B x C x D	2	.11	.05	.19	

Source	df	SS	MS	F	p
A x B x C x D	4	1.56	.39	1.36	
Error	108	31.04	.29		

Appendix 9

Summary of Analysis of Variance of Scores

on Allocated Punishment

(Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	8.47	4.24	5.14	<.01
B: Cause	2	5.98	2.99	3.63	<.03
A x B	4	10.61	2.65	3.22	<.02
Error	108	88.92	.82		
C: Severity	1	55.54	55.54	141.30	<.001
A x C	2	.23	.12	.29	
B x C	2	.72	.36	.91	
A x B x C	4	1.05	.26	.67	
Error	108	42.46	.39		

Appendix 10

Summary of Analysis of Variance of Scores on Allocated Reward (Experiment 1)

Source	df	SS	MS	F	p
A: Grades	2	8.78	4.39	2.79	<.07
B: Cause	2	15.29	7.65	4.86	<.01
A x B	4	18.33	4.58	2.91	<.03
Error	108	169.77	1.57		
C: Severity	1	22.77	22.77	41.47	<.001
A x C	2	.26	.13	.24	
B x C	2	.78	.39	.71	
A x B x C	4	2.38	.59	1.08	
Error	108	59.31	.55		

Appendix 11

Description of Stimulus Materials

for Experiment 2

Bad-Outcome/High-Severe Story

Facilitative-cause condition. "Almost all of the kids in Bobby's school have sometimes taken bicycles that didn't belong to them, and kept them for themselves. Bobby took another boy's bicycle and kept it for himself."

Inhibitory-cause condition. "Almost none of the kids in Bobby's school have ever taken bicycles that didn't belong to them. Bobby took another boy's bicycle and kept it for himself."

Control condition. "Bobby took another boy's bicycle and kept it for himself."

Dependent measures. A. "Afterwards, how responsible was Bobby for the other boy not having the bicycle anymore?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How much should Bobby be punished for taking the other boy's bicycle?" (1) not punished at all, (2) punished a little bit, (3) punished pretty much, (4) punished a lot. C. "How often does Bobby take other things besides bicycles that don't belong to him?" (1) not often at all, (2) once in awhile, (3) pretty often, (4) almost all of the time. D. "How many of the kids in Bobby's school would have taken the other boy's bicycle if they had the chance?" (1) almost none of the kids, (2) a few of the kids, (3) most of the kids, (4) almost all of the kids.

Appendix 11 (continued)

Bad-Outcome/Low-Severe Story

Facilitative-cause story. "Almost all of the kids in John's school have sometimes taken pencils that didn't belong to them, and kept them for themselves. John took another boy's pencil and kept it for himself."

Inhibitory-cause story. "Almost none of the kids in John's school have ever taken pencils that didn't belong to them. John took another boy's pencil and kept it for himself."

Control condition. "John took another boy's pencil and kept it for himself."

Dependent measures. A. "Afterwards, how responsible was John for the other boy not having the pencil anymore?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How much should John be punished for taking the other boy's pencil?" (1) not punished at all, (2) punished a little bit, (3) punished pretty much, (4) punished a lot. C. "How often does John take other things besides pencils that don't belong to him?" (1) not often at all, (2) once in awhile, (3) pretty often, (4) almost all of the time. D. "How many of the kids in John's school would have taken the other boy's pencil if they had the chance?" (1) almost none of the kids, (2) a few of the kids, (3) most of the kids, (4) almost all of the kids.

Good-Outcome/High-Severe Story

Facilitative-cause condition. "Almost all of the kids in

Appendix 11 (continued)

Steve's school have sometime given all of their savings, which is a lot of money, to poor children. Steve gave all of his savings, which was also a lot of money, to poor children."

Inhibitory-cause condition. "Almost none of the kids in Steve's school have ever given any of their savings to poor children. Steve gave all of his savings, which was a lot of money, to poor children."

Control Condition. "Steve gave all of his savings, which was a lot of money, to poor children."

Dependent measures. A. "Afterwards, how responsible was Steve for the poor children having all of his savings?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How much should Steve be rewarded for giving his savings to poor children?" (1) not rewarded at all, (2) rewarded a little bit, (3) rewarded pretty much, (4) rewarded a lot. C. "How often does Steve help other people in other ways besides giving them money?" (1) not often at all, (2) once in awhile, (3) pretty often, (4) almost all of the time. D. "How many of the kids in Steve's school would have given their savings to poor children if they had the chance?" (1) almost none of the kids, (2) a few of the kids, (3) most of the kids, (4) almost all of the kids.

Appendix 11 (continued)

Good-Outcome/Low-Severe Story

Facilitative-cause condition. "Almost all of the kids in Andy's school have sometimes helped younger children tie their shoelaces. Andy helped a little boy on the playground tie his shoelaces."

Inhibitory-cause condition. "Almost none of the kids in Andy's school have ever helped younger children tie their shoelaces. Andy helped a little boy on the playground tie his shoelaces."

Control condition. "Andy helped a little boy on the playground tie his shoelaces."

Dependent measures. A. "Afterwards, how responsible was Andy for the little boy having his shoelaces tied?" (1) not at all responsible, (2) a little bit responsible, (3) pretty responsible, (4) very responsible. B. "How much should Andy be rewarded for tying the little boy's shoelaces?" (1) not rewarded at all, (2) rewarded a little bit, (3) rewarded pretty much, (4) rewarded a lot. C. "How often does Andy help other people in other ways besides tying shoelaces?" (1) not often at all, (2) once in awhile, (3) pretty often, (4) almost all of the time. D. "How many of the kids in Andy's school would have helped the little boy tie his shoelaces if they had the chance?" (1) almost none of the kids, (2) a few of the kids, (3) most of the kids, (4) almost all of the kids.

Appendix 12

Summary of Analysis of Variance of Scores

on Judgments of Consensus

(Experiment 2)

Source	df	SS	MS	F	p
A: Grades	2	.96	.48	.38	
B: Cause	2	166.86	83.43	66.79	<.001
A x B	4	3.22	.81	.64	
Error	117	146.16	1.25		
C: Valence	1	6.00	6.00	6.73	<.02
A x C	2	42.22	21.11	23.66	<.001
B x C	2	3.38	1.69	1.90	
A x B x C	4	6.27	1.57	1.76	
Error	117	104.37	.89		
D: Severity	1	3.02	3.02	9.31	<.01
A x D	2	9.37	4.68	14.44	<.001
B x D	2	3.73	1.86	5.74	<.01
A x B x D	4	2.19	.55	1.69	
Error	117	37.95	.32		
C x D	1	3.34	3.34	10.08	<.01
A x C x D	2	.79	.39	1.19	
B x C x D	2	.36	.18	.55	

Source	df	SS	MS	F	p
A x B x C x D	4	1.03	.26	.78	
Error	117	30.73	.33		

Appendix 13

Summary of Analysis of Variance of Scores on Attributions of Responsibility (Experiment 2)

Source	df	SS	MS	F	p
A: Grades	2	11.46	5.73	3.79	<.05
B: Cause	2	2.34	1.17	.77	
A x B	4	1.97	.49	.33	
Error	117	176.89	1.51		
C: Valence	1	10.87	10.87	10.58	<.002
A x C	2	.98	.49	.48	
B x C	2	5.08	2.54	2.47	<.09
A x B x C	4	3.40	.85	.83	
Error	117	120.18	1.03		
D: Severity	1	5.79	5.79	11.56	<.001
A x D	2	.01	.005	.01	
B x D	2	.94	.47	.94	
A x B x D	4	5.23	1.31	2.61	<.05
Error	117	58.54	.50		
C x D	1	.79	.79	1.80	
A x C x D	2	.60	.30	.68	
B x C x D	2	.36	.18	.41	

Source	df	SS	MS	F	p
A x B x C x D	4	1.21	.30	.69	
Error	117	51.54	.44		

Appendix 14

Summary of Analysis of Variance of Scores on Attributions of Diagnosticity (Experiment 2)

Source	df	SS	MS	F	p
A: Grades	2	2.98	1.49	1.19	
B: Cause	2	2.50	1.25	1.00	
A x B	4	3.57	.89	.71	
Error	117	146.62	1.25		
C: Valence	1	41.72	41.72	66.94	<.001
A x C	2	6.86	3.43	5.51	<.01
B x C	2	2.86	1.43	2.30	
A x B x C	4	2.90	.73	1.16	
Error	117	72.91	.62		
D: Severity	1	.05	.05	.12	
A x D	2	1.98	.99	2.44	<.10
B x D	2	1.93	.97	2.39	<.10
A x B x D	4	1.88	.47	1.16	
Error	117	47.41	.41		
C x D	1	1.24	1.24	3.39	<.07
A x C x D	2	1.43	.72	1.96	
B x C x D	2	.15	.07	.20	

Source	df	SS	MS	F	p
A x B x C x D	4	3.59	.90	2.45	<.05
Error	117	42.84	.37		

Appendix 15

Summary of Analysis of Variance of Scores on Allocated Punishment (Experiment 2)

Source	df	SS	MS	F	p
A: Grades	2	11.17	5.58	9.97	<.001
B: Cause	2	1.81	.90	1.62	
A x B	4	2.67	.67	1.19	
Error	117	65.54	.56		
C: Severity	1	83.43	83.43	246.90	<.001
A x C	2	3.63	1.81	5.37	<.01
B x C	2	.32	.16	.05	
A x B x C	4	.87	.22	.65	
Error	117	39.54	.34		

Appendix 16

Summary of Analysis of Variance of Scores on Allocated Reward (Experiment 2)

Source	df	SS	MS	F	p
A: Grades	2	17.17	8.58	5.32	<.01
B: Cause	2	1.17	.58	.36	
A x B	4	5.67	1.42	.88	
Error	117	188.71	1.61		
C: Severity	1	26.68	26.68	47.30	<.001
A x C	2	3.96	1.98	3.51	<.05
B x C	2	2.06	1.03	1.82	
A x B x C	4	2.30	.58	1.02	
Error	117	66.00	.56		

Appendix 17

Intercorrelations of Major Dependent Variables of Experiment 1

	Bad Outcomes		
	Diagnosticity	How Bad	Punishment
		Grade 2	
Responsibility	.00	.20*	.24*
Diagnosticity		.14	.14
How Bad			.31**
		Grade 5	
Responsibility	.29**	.59**	.39**
Diagnosticity		.35**	.35**
How Bad			.37**
		Grade 8	
Responsibility	-.02	.24*	.31**
Diagnosticity		.01	.05
How Bad			.42**

* $p < .05$

** $p < .01$

n = 78

Appendix 17 (continued)

	Good Outcomes		
	Diagnosticity	How Good	Reward
		Grade 2	
Responsibility	-.06	.08	.13
Diagnosticity		.20*	.28**
How Good			.65**
		Grade 5	
Responsibility	.30**	.28**	.15
Diagnosticity		.34**	.29**
How Good			.47**
		Grade 8	
Responsibility	.21*	.31**	.20*
Diagnosticity		.26*	.36**
How Good			.21*

* $p < .05$

** $p < .01$

n = 78

Appendix 18

Intercorrelations of Major Dependent Variables of Experiment 2

Bad Outcomes

Grade	Responsibility/ Punishment	Responsibility/ Diagnosticity	Punishment/ Diagnosticity
2	.42**	.03	.15
5	.30**	-.05	.12
8	.22*	.01	.36**

Good Outcomes

Grade	Responsibility/ Reward	Responsibility/ Diagnosticity	Reward/ Diagnosticity
2	.11	.06	.27**
5	.39**	.32**	.29**
8	.21*	.16	.49**

* $p < .05$

** $p < .01$

n = 84